THE UNIVERSITY OF ALBERTA

COMPREHENSION AND PRODUCTION OF WORD ORDER BY TWO-YEAR OLD CHILDREN

by

Janet Adele McClellan

(C)

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES AND RESEARCH

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE

OF MASTER OF EDUCATION

DEPARTMENT OF EDUCATIONAL PSYCHOLOGY

EDMONTON, ALBERTA

FALL, 1979

THE UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES AND RESEARCH

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research, for acceptence, a thesis entitled Comprehension and Production of Word Order by Two-Year Old Children submitted by Janet Adele McClellan in partial fulfilment of the requirements for the degree of Master of Education

Supervisor

Date Oct 18, 1979

ABSTRACT

This study is an an attempt to test the predictions made by Chapman and Miller (1975) that (1) productive control of the subject-object relation exceeds comprehension performance based on syntactic form alone, and (2) less linguistically advanced children decode subject-object relations lacking referential support on the basis of a lexical-semantic strategy which assigns animate nouns to subject status and inanimate nouns to object status. Ten children, divided on the basis of mean length of utterance into two groups (average MLU of 2.45 and 3.75, respectively), participated. Comprehension and production of subject-object order in semantically reversible sentences with inanimate or animate subject and object was assessed using an object manipulation format.

The results obtained provided partial support for the predictions made by Chapman and Miller (1975). The proposed lexical-semantic strategy for comprehension was not employed by the subjects of the present study; these children appeared to use a probable event strategy in determining subject-object assignment. Although productive control of the ordering of subject and object was found to exceed comprehension performance, the facilitating effect of visual cues available during the production trials likely contributed to the observed differences in performance between the two tasks.

ACKNOWLEDGMENTS

I would like to extend sincere thanks to Dr. Carolyn Yewchuk for her supervision and support throughout all phases of this study. Gratitude is also expressed to Dr. G. Holdgrafer, Dr. Gene Romaniuk and Dr. W.H.O. Schmidt for their discussion and careful reading of the manuscript.

The mothers who welcomed me into their homes must also be warmly acknowledged. In addition, thanks is extended to Mrs. Shirley Stawnychy who typed the final draft.

Most of all, I thank the children, Mark, Sara, Katrina, Kerry, Michael, Nathan, Chantelle, Levi, Shannon and Anna-Marie. It was their co-operation (and occasional resistance) that made this project an enjoyable and worthwhile undertaking.

Finally, to Jim, my appreciation and love for his patience and encouragement during the past year.

TABLE OF CONTENTS

CHAPTER		PAGE
1	INTRODUCTION	1
2	REVIEW OF THE LITERATURE	4
	Perspective	4
	Studies: Comprehension Versus Production of Sentence Structure	8
	Theoretical Views: The Relationship Between Comprehension and Production	22
	Ingram (1974)	22
	Bloom (1974;1978)	26
	Summary	29
	A Detailed Review: Chapman and Miller (1975).	29
	Critical Analysis of Chapman and Miller (1975).	37
3	HYPOTHESES AND DEFINITIONS	45
	Hypotheses	45
	Definitions	46
4	EXPERIMENTAL PROCEDURE AND DESIGN	48
	Subjects and Groups	50
	Speech Sample	51
	Sentences	52
	Stimulus Objects	54
	Comprehension Task	54
	Production Task	55

TABLE OF CONTENTS, CONTINUED

CHAPTER		PAGE
4	Pretesting	55
	Comprehension Scoring	56
	Production Scoring	57
	Procedures	57
	Reliability Measures	58
5	RESULTS	60
	Three-Way Analysis of Variance of Error Scores	
	Main Factors	60
	Interactions	60
	Post-hoc Contrasts	63
	Two-Way Analysis of Variance of Scorable Responses	66
	Main Factors	66
	Interaction	67
6	DISCUSSION	68
	Hypothesis I	68
	Hypothesis II	69
	Hypothesis III	79
	Hypothesis IV	82
7	CONCLUSIONS, IMPLICATIONS and LIMITATIONS	86
	BIBLIOGRAPHY	93

LIST OF TABLES

TABLE		PAGE
1	Summary of the experimental literature	20
2	Chapman and Miller (1975): data summary	31
3	Chapman and Miller (1975): stimulus sentences	32
4	Chapman and Miller (1975): scoring criteria	34
5	Chapman and Miller (1975): Percent correct for sentences scorable on both tasks	36
6	Chapman and Miller (1975): comprehension of sentence types: mean percent correct by group	36
7	Chapman and Miller (1975): production of sentence types: mean percent by group	36
8	Chapman and Miller (1975): mean difference in number of scorable responses between tasks	38
9	Chapman and Miller (1975): scorable versus correct responses	39
10	Sentence types	46
11	Sex, age and MLU values for all groups	51

LIST OF TABLES, CONTINUED

TABLE		PAGE
12	Test sentences for the comprehension and production tasks	53
13	Data summary	61
14	Three-way analysis of variance: main and interaction effects	62
15	Post-hoc contrast between tasks for given groups	64
16	Post-hoc contrast between tasks for given sentence types	64
17	Post-hoc contrast among sentence types for given tasks	66
18	Two-way analysis of variance: main and interaction effects	66
19	Mean percent of errors scored on sentence types for Chapman and Miller (1975) and the present study	79

LIST OF FIGURES

FIGURE		PAGE
1	Interaction between tasks and sentence types	72
2	Comprehension independent of context	75
3	Production with contextual support	75

CHAPTER ONE

INTRODUCTION

The traditional view of the relationship between comprehension and production assumed by McCarthy (1954) and supported by Ingram (1974) has recently been challenged in the literature. In particular, Bloom (1978) and Clark et al. (1974) have called for the re-examination of a tenet that has become axiomatic in theories of language acquisition. Specifically in question is the hypothesis that language comprehension precedes production in ontogeny. Growing skepticism regarding the validity and utility of this position has been nurtured to some extent by studies which provide evidence either unaccounted for or in direct opposition to this long unquestioned precept.

Such evidence is typified in the work of Chapman and Miller (1975). These authors challenge the hypothesis that comprehension of syntactic form, specifically, word order, precedes production. In their study, children whose productive utterances ranged in length from 1.53 to 3.11 morphemes were tested on both comprehension and production tasks in which attention to appropriate word order was crucial to successful performance. The ability of the child to comprehend a sentence on the basis of word order alone was compared to her ability to correctly encode a specific grammatic relation (subject-object) in a production task. The test sentences were active, reversible and

constructed from the use of three animate, three inanimate and six transitive verbs. Four sentence types were formed on the basis of the animacy-inanimacy of the subject and object.

Results indicated that the children were successful in preserving the appropriate subject-verb-object ordering in the production task significantly more often than they were able to use word order as a strategy leading to fruitful performance in the comprehension task. From this, Chapman and Miller (1975) concluded that

> "the competence to be attributed to the child on the basis of comprehension is less advanced and different from the competency to be attributed on the basis of production for the ordering of subject and object" (p. 367).

In addition, it was found that comprehension performance varied as a function of sentence type; however, similar patterns of performance were not evident for the production task. From this the authors concluded that children employ different strategies when comprehending and producing sentences. For comprehension, Chapman and Miller (1975) propose that less linguistically advanced children decode subject-object relations on the basis of a lexical-semantic strategy which assigns animate nouns to subject status and inanimate nouns to object status.

Although widely quoted in the literature, it is believed here that the conclusions reached by Chapman and Miller (1975) should be considered with caution. Reticence in accepting this study as evidence counter to the

traditional position arises from (1) methodological flaws inherent in their design, (2) the questionable statistical analyses of the data employed and (3) the availability of alternate interpretations of the data.

The present study introduces changes in procedures and analysis in an attempt to test the prediction made by Chapman and Miller (1975) that the child's productive control of the subject-object relation will exceed comprehension performance based on syntactic form alone. Further, this analysis will incidentally provide data that may be used to verify the lexical-semantic strategy for comprehension proposed by these authors.

CHAPTER TWO

REVIEW OF THE LITERATURE

Perspective

The work of Braine (1963), Brown and Fraser (1964) and Miller and Ervin (1964) reflects the modern grammatical approach to child language, an orientation which saw heuristics gained from American structural and transformational linguistics employed in the analysis of early, multiword utterances. Rather than considering child language as simply a reduced form of the adult code, these investigators sought to formulate a grammar peculiar to the language spoken by children in the process of development. Such initial attempts have been subsequently criticized on various grounds by Bloom (1970), Bowerman (1973) and Brown (1973) and the seventies have witnessed a general trend away from purely syntactic models to paradigms that hypothesize a cognitive-semantic grammatic base. (Fillmore, 1968; Slobin, 1970; Schlesinger, 1971; Brown, 1973; Bates, 1976).

Whether semantically or syntactically disposed, the grammatic studies have chosen multiword utterances as their primary datum. An alternate emphasis, one which is apparent in earlier systematic accounts of language acquisition (Guillaume, 1927; DeLaguna, 1927; Stern and Stern, 1930; Leopold, 1949) has been revitalised recently in the work of Dore (1975), Halliday (1975), Bates (1976) and Greenfield and Smith (1976). These investigators have devoted

careful attention to the period of single word utterances, commonly referred to as the "holophrastic" stage. Several theoretical issues identified in the course of these studies have become topics of lively debate. In particular, the existence of relational or combinatorial meaning prior to the stage of multiword utterances is a subject that is yet to be fully resolved.

Greenfield and Smith (1976) argue strongly for the relational status of single word utterances and do so from a grammatic framework that is semantically based. It is their contention that the child using only single words intends relational meaning and achieves its expression through combining the word with relevant components of the nonlinguistic context.

In contrast, Bloom (1973) insists that the child cannot be reliably credited with intending relational meaning unless its expression is realized through word combination.

A detailed analysis of the arguments and empirical evidence provided by these authors will not be pursued here. However, of special interest are the views of each regarding the presyntactic child's knowledge of sentence structure. Specifically at issue is the extent to which the child, whose productive ability is restricted to single words, utilizes syntactic information in comprehending speech.

Greenfield and Smith (1976) cite evidence from Huttenlocker (1974) to support the notion that the child

who is restricted to producing single word utterances has a richer underlying conceptual structure to which the word can be related. The evidence is found in the ability of two children at the single word stage to respond relationally to information presented in adult speech. For example, one child was able to differentially respond to sentences like "Show the baby's bottle to Mommy" and "Show Craig's bottle to the baby". According to Greenfield and Smith (1976), this required the child to assign "baby" the roles of possessor and recipient and further that

> "to do so required syntactic processing, at the very least processing of information about word order - thus, a stage of syntactic comprehension occurs in the one word stage but follows a stage in which the child relates individual words to the situational structure as he perceives it" (p. 220).

In sum, the authors posit that even prior to the stage of word combination, the child acquires specifically linguistic knowledge and uses this information in the process of language comprehension.

A view radically opposed to this analysis is found in Bloom (1973;1978). The assumption that the child's comprehension exceeds productive ability at every stage of language development and its corollary that the analysis of sentence structure (syntax) is somehow involved in such comprehension, are refuted by Bloom. Instead, she proposes that children possess little knowledge of linguistic structure prior to the use of syntax as evidenced through word combination. At the single word stage, the child's

appropriate responses to statements and directions are believed to be a function of the redundancy of the linguistic message with extralinguistic information (gesture, intonation, exaggeration) provided by the immediate context. Simply, the child need not analyse the structure of such utterances in order to respond appropriately; the necessary cues are provided within the communicative situation. Bloom (1973) explains that

> "when a sentence is redundant with respect to the context in which it occurs, then the amount of information which the child needs to get from the linguistic message is probably minimal...it remains to be tested whether children using single word utterances understand sentences in non-redundant contexts as well as they understand what is said to them about what they can see or hear or do ... if they do not understand sentences that refer to relations among objects and events that are not immediately available, then the extent to which they analyse syntactic structure in their understanding is certainly questionable" (p. 56-59).

In summary, Greenfield and Smith (1976) believe that processing of syntactic information (word order) as a primary aid to comprehension occurs prior to the stage in which words are combined in productive speech. Alternately, Bloom (1973;1978) contends that comprehension of sentences by the presyntactic child depends exclusively on lexical and contextual cues. The child is attributed with no syntactic knowledge.

The position of each regarding the nature of the relationship between comprehension and production of syntactic form may be inferred from the above comments. Where Greenfield and Smith (1976) hold that syntactic comprehension

precedes production, Bloom (1973;1978) proposes that the understanding of linguistic structure need not precede its production but may evolve as the child discovers the various means of word combination.

Several studies addressing this issue will be reviewed in the following section. These have employed as subjects children functioning both at the single word stage and beyond.

Studies: Comprehension Versus Production of Sentence Structure

Fraser, Bellugi and Brown (1963) tested the familiar assertion that language comprehension precedes production in a study in which the ability of twelve children aged between 37 and 43 months to imitate, comprehend and produce ten grammatic contrasts was assessed. Two pictures illustrating a grammatic contrast (eg. The daddy kisses the mommy; the mommy kisses the daddy) were employed as stimulus items. Comprehension was operationalized as the correct identification via selective pointing to pictures named by contrasting sentences. Production was defined as either the correct imitation of contrasting features in sentences or the correct production of contrasting features in sentences applied appropriately to pictures. The procedural format for their study has since been termed the ICP Test. Results indicated that production as defined in the latter case proved to be less advanced than understanding for their subjects. Thus, the thesis that comprehension exceeds productive control is supported.

Replications of the basic procedures employed by Fraser et al. (1963) appear in the literature. Lovell and Dixon (1967) tested one hundred children ranging in age from two to six years on the ten grammatic contrasts taken, with slight amendment, from the ICP Test. In accordance with the findings of Fraser et al. (1963), imitation was more advanced than comprehension and the latter exceeded production for all subjects and age levels. Using similar procedures, Nurss and Day (1971) assessed whether the same sequence of development in the control of grammar as outlined in the previous studies occurred for higher status white and lower status black and white four year olds. The imitation-comprehension-production sequence was supported in part by their data. While performance on both the comprehension and imitation tasks exceeded that of the production task, no significant difference was found between the subjects' ability to imitate and comprehend the grammatic contrasts. Cohen (1967) and Turner and Rommetveit (1967) provide further empirical support for the I-C-P sequence with older subjects.

However, the Fraser et al. (1963) study and its replications have not escaped scrutiny. In fact, papers by Fernald (1972) and Baird (1972) have cast serious doubt on the validity of the methodology and scoring procedure used in these investigations.

Baird (1972) demonstrated that the direct comparison of scores obtained on the ICP tasks was inappropriate since

the necessary condition for their comparison, that is, equal or controlled probabilities of chance success for each task, was not met. Thus, the obtained differences resulting from comparisons of task performance as made by Fraser et al. (1963) and others, may reflect different chance success rates rather than true differences in the difficulty of each task. Fernald (1972) offers a similar criticism by identifying the control problem inherent in the ICP scoring procedure. Specifically, in the Fraser et al. (1963) scoring procedure, novel responses (ie. sentences that were gramatically correct but not the specific sentence required) were scored as errors. According to Fernald (1972), this procedure is problematic in that it tends to artificially inflate the number of errors on the production task compared to the comprehension task. Since the comprehension task required the choice of one of two pictures, the probability of a chance success equals 0.50. However, the production task offered a wider possibility of responses resulting in a chance success probability of less than 0.50. Thus, the unequal probabilities of chance success inherent in the scoring procedure biased the scores in favor of comprehension whenever the child guessed.

Fernald (1972) hypothesized that the results obtained in the ICP studies were an artifact of the scoring procedure rather than a reflection of comprehension or production ability. To test his hypothesis, a replication of Fraser et al. (1963) was conducted with the differential probabil-

ities of chance success being statistically controlled. It was found that when the scoring artifact was removed, no significant difference in performance of the two tasks existed.

The studies examined to this point have used as subjects children three years of age and older. Although mean length of utterance (MLU) data are not reported, it may be assumed that the children tested were well into the period of word combination. The studies which follow have reported MLU values and have included subjects whose productive ability is confined basically to single word utterances. In addition, where the earlier studies used a picture format for both tasks, subsequent investigations have employed an action format in their procedures.

In an attempt to discover whether a child's spontaneous utterances can be taken as a direct indication of her linguistic competence, Shipley, Smith and Gleitman (1969) conducted a study in which spontaneous action responses to commands differing in structural complexity and semantic content were assessed for two groups of young speakers. Children were grouped on the basis of MLU into either a holophrastic (1.06-1.16 morphemes) or telegraphic (1.40-1.85 morphemes) group.

The plan of the experiment was to deliver commands varied systematically in syntax and content to young speakers in order to determine whether differential responding to varied constructional types would occur. The

authors assumed that different responses to stimuli whose semantic content was identical but whose syntactic structure differed may be taken as evidence that the child noticed and was affected by such differences. The stimulus categories included four forms: (1) well formed command eg. Throw me the ball; (2) telegraphic command eg. Throw ball; (3) lengthened telegraphic command eg. Please John, throw ball; and, (4) noun only command eg. Ball!. After each command, the child's action and/or verbal response was recorded. To be credited with a correct response, the child had to at least touch the object named in the command.

Results indicated that the telegraphic group obeyed well formed commands significantly more often than those of categories (2), (3) and (4). An opposing trend was apparent for the holophrastic group who tended to obey the child form commands of categories (2), (3) and (4) more frequently than the well formed stimuli. From this, the authors conclude that where the holophrastic child <u>prefers</u> to respond to speech just at or slightly above her productive ability, the telegraphic speaker responds best to just those utterance types she does not use. Further, the telegraphic child is capable of discriminating more speech forms than she is able to use ie. syntactic comprehension precedes production for the telegraphic speaker. This implies that attention to sentence form becomes more significant in comprehension when the child moves from the stage of single words to the period of word combination.

Recently, Petretic and Tweney (1977) replicated with modifications the basic procedures employed in the Shipley et al. (1969) experiment. These authors distinguished several methodological flaws inherent in the original design. First, the scoring procedure did not allow for accurate determination of comprehension since it included a wide range of diverse behaviors. Second, while group averages for subjects in the holophrastic group indicated a trend in favor of child form commands, a consistent pattern of responding was not apparent since two of the four subjects responded with equal or greater frequency to the well formed commands. Finally, that mothers were employed to deliver test stimuli may have introduced confounding factors due to lack of uniformity of presentation.

Several procedural modifications were introduced by Petretic and Tweney (1977) to control for these factors including (1) a greater number of subjects (2) well defined response categories and, (3) the use of a single experimenter to deliver stimuli whose familiarity with all subjects remained constant.

Consistent with the Shipley et al. (1969) results, the authors found that advanced telegraphic speakers responded most frequently to adult form commands. However, contrary to the earlier findings, all children, even the earliest telegraphic speakers (referred to as holophrastic by Shipley et al., 1969) were more accurate in responding to adult forms than child forms. The authors suggest that

the ability to comprehend adult linguistic forms is present even at the earliest stages of word combination.

Wetstone and Friedlander (1973) designed an experiment to investigate how specific elements of syntax affect the young child's sentence comprehension. Working under the assumption that if word order carries communicative value its distortion should disrupt the effectiveness of a question or command, the authors tested 20 children who were placed on the basis of MLU into three groups: holophrastic (mean MLU of 1.75), telegraphic (mean MLU of 2.79) and fluent (mean MLU of 3.73). Each child was required to carry out a number of questions and commands that were presented to them in sentences possessing varying degrees of word order distortion eg. Open up the box; Box the up open; Open the up box. The child's appropriate response to each command was taken as an index of its communicative effectiveness.

Results indicated that for the holophrastic and telegraphic groups, no response score differences existed between the well formed and scrambled sentences. Alternately, the fluent group presented significantly lower scores for the scrambled sentences compared to those of normal word order. From this the authors conclude that word order carries little or no communicative value for the nonfluent child. Further, they suggest that at this stage, receptive language processing is focused on familiar semantic elements and immediate extralinguistic context rather than syntactic

information. However, this did not seem to be the case for the fluent speaker. Since the fluent group responded appropriately less frequently to the sentences with distorted word order, the authors infer that both semantic and syntactic information is significant for comprehension at this stage.

Methodological weakness apparent in the Wetstone and Friedlander (1973) design signals the need for a re-examination of their findings. Most obviously, the test sentences were such that attention to word order was not critical for the correct interpretation of the sentence. For example, for the test sentence "Show me how you jump", the child need only attend to the word "jump" to respond appropriately since the carrier phrase "show me how you" is redundant with the context and carries limited communicative value. In contrast, a reversible sentence such as "Show me how the boy hits the girl" is critically dependent on word order for correct interpretation. None of the sentences used by Wetstone and Friedlander (1973) were of this type.

Secondly, the sentences were presented in redundant contexts with normal intonation and no precautions are mentioned to ensure that the actions required by the test sentences did not constitute some of the child's established routine. In effect, the authors' results may have been confounded by extralinguistic factors operative in the testing environment. Thirdly, the fact that the overall level of performance for all groups was high, with the mean

number of relevant responses to normal, misplaced, and scrambled word order being 90, 81.5 and 79.5 percent, respectively, may also indicate the dubious quality of the test sentences as reliable, discriminating variables.

Finally, the significant difference in performance noted between the nonfluent (holophrastic and telegraphic) and fluent groups may be an artifact of the scoring procedure used. Included as incorrect responses were irrelevent or out of context responses, failures to respond and indications of puzzlement or demands for explanation. It is not indicated what proportion of responses classed as incorrect were of each type and it may be that none of the above are an absolute indication that the child misinterpreted the message due to a reliance or nonreliance on word order. For example, the younger children of the nonfluent group may not have been as generally sophisticated as to be able to request a repetition or demand an explanation.

In sum, it is uncertain on the basis of this study whether children who are presyntactic or in the very early stages of word combination are able to use word order information when it is critical to the interpretation of a message.

DeVilliers and DeVilliers (1973) investigated the young child's use of word order in a study methodologically superior to that of Wetstone and Friedlander (1973). Thirty-three children divided on the basis of MLU (ranging from 1.06 to 4.25 morphemes) into six groups were presented

with six reversible active sentences and required to act them out eg. Make the cat bite the dog. A comparison of the results of the two studies proves interesting. In the DeVilliers (1973) experiment, only the seven subjects of the least linguistically mature group (MLU of 1.00 to 1.50) were unable to use word order in the active reversible sentence comprehension task. Seventeen linguistically more advanced subjects (MLU of 1.50 to 3.00) were able to perform actions in response to the test sentences an average of 75 percent of trials. Thus, contrary to Wetstone and Friedlander (1973), the subjects in this experiment who were approximately equivalent in MLU to Wetstone and Friedlander's telegraphic group did use word order as a comprehension strategy with reasonable accuracy.

Chapman and Miller (1975) performed a study in which the child's ability to correctly encode a subject-verb-object relation was compared to her ability to use word order as an aid to comprehension. A full examination of this study will be pursued in a subsequent section. Briefly, their results indicated the children tested scored significantly better on the production task than on the comprehension task. In effect, production was found to exceed comprehension for the subject-object relation.

Keeny and Wolfe (1972) tested the production, imitation and comprehension of subject-verb agreement for number (eg. The bird is singing; the birds are singing) in sentences with 46 three and four year olds. Spontaneous speech samples

indicated the subjects reliably inflected verbs for number in naturalistic settings with 94 percent accuracy. In the imitation task, sentence stimuli were either grammatical or ungrammatical for subject-verb agreement. Verbal responses were scored with 83 percent of grammatic sentences being repeated verbatim compared to 93 percent of ungrammatic sentences repeated not verbatim but in corrected form. Poor performance on the comprehension subtests lead the authors to conclude that the subjects did not understand the relation between verb inflections for number and the meaning of singular versus plural. However, since subject-verb agreement was correctly produced in speech, Keeny and Wolfe (1972) concluded that production precedes comprehension for subject-verb number agreement.

Lahey (1974) conducted a study which indicated children may incorporate surface features of the language into their speech prior to their understanding of the underlying meaning distinctions signalled by such forms. She found that in interpreting sentences such as "The cow that hit the pig chased the sheep", four and five year old subjects appeared to erroneously rely on a word order strategy rather than using syntactic markers as a cue to sentence meaning. It is notable that children at this age are reported to produce several syntactic markers in their spontaneous speech.

The studies presented in this section have attempted to determine the nature of the relationship between comprehension and production of sentence structure for young

speakers. It may be useful to view in summary form the conclusions of each found in Table 1.

The information provided thus far allows two general conclusions. First, it remains uncertain whether the comprehension of sentence structure precedes the child's ability to encode specific grammatic relations through the use of syntax. Second, the young child may employ any combination of semantic, syntactic, nonlinguistic or contextual cues or strategies in her attempt to comprehend speech. Further, the relative contribution of each may shift depending on the level of linguistic development or the particular child.

The lack of consensus outlined in these general conclusions may be attributed to several factors. First, the studies examined have differed in the populations, methodologies and scoring-analysis procedures used. Consider, for example, the action versus the picture format, the variability in MLU limits for assignment to linguistic group, the lack of uniform scoring criteria and the mutable environmental control procedures. In effect, the different results obtained by the authors listed may be attributable to such variations. Secondly, the children used as subjects generally ranged in age from twelve to thirty-six months. This population presents several methodological challenges since the young child is easily distracted and frequently uncooperative. Finally, accurate measurement of the young child's sentence comprehension remains somewhat of a methodological enigma.

Table 1. Summary of Experimental Literature

	Study	Conclusions
1.	Fraser et al. (1963)	Imitation exceeds comprehension which exceeds production (I-C-P) for ten grammatic contrasts.
2.	Lovell and Dixon (1967)	Verification of the I-C-P sequence.
3.	Nurss and Day (1971)	Partial verification of the I-C-P sequence with no significant difference between I and C.
4.	Cohen (1967); Turner and Rommetveit (1967)	Verification of the I-C-P sequence.
5.	Baird (1972)	Scoring procedure used in (1), (2), (3) and (4) shown to be biased in favor of comprehension.
6.	Fernald (1972)	Replication of (1) with scoring bias controlled found no significant difference between C and P performance.
7.	Shipley et al. (1969)	Comprehension of syntax precedes production for the telegraphic speaker.
8.	Petretic and Tweney (1977)	Replication of (7) with procedural improvement found comprehension of sentence form in advance of production for both holophrastic and telegraphic speakers.
9.	Wetstone and Friedlander (1973)	Syntactic cues not used in sentence comprehension by holophrastic and telegraphic speakers.
10.	DeVilliers and Devilliers (1973)	In a comprehension task, word order cues were used successfully by telegraphic but not holophrastic subjects.

Table 1, continued

11.	Chapman and Miller (1975)	Production precedes comprehension for the ordering of subject-object.
12.	Keeny and Wolfe (1972)	Production exceeds comprehension for subject-verb number agreement.
13.	Lahey (1974)	Production of surface features may precede the comprehension of the distinctions signalled by such forms.

The confounding factors specific to the analysis and assessment of sentence comprehension will receive attention in subsequent sections. What is important to emphasize here, however, is that the procedural problems encountered in studies of early sentence comprehension are directly related to the complex nature of the process itself. To clarify, the listener uses semantic, syntactic, paralinguistic and contextual information when attempting to decode speech. In effect, strategies of various types are employed. Thus, what is crucial to the analysis of sentence comprehension are operational definitions and methods that are of sufficient sophistication to allow for the effective control of potentially confounding variables.

The following section will further underline this necessity. It examines the nature and relationship of comprehension and production in theory.

Theoretical Views: The Relationship Between Comprehension and Production

Ingram (1974): The Traditional View

Ingram (1974) argues from the traditional standpoint that the relationship between comprehension and production is a unidirectional one in which some comprehension of a specific construction must precede its production. He states

> "comprehension does precede production... it could never be any other way... it is proposed that comprehension ahead of production is a linguistic universal of acquisition" (p. 313).

According to Ingram, proposed counterevidence (Bloom, 1974; Chapman and Miller, 1975) in favor of alternate hypotheses regarding this issue is based in many cases upon a misinterpretation of the traditional position. Further, the empirical issues raised by such opposing studies are thought to involve the nature of comprehension and production and the gap between them rather than whether one precedes the other. In sum, Ingram proposes that research involving comprehension and production has produced considerable confusion due to methodological and definitional inconsistency.

Ingram (1974) attempts to clarify the traditional position by delineating its common misinterpretations. In addition, he attempts to demonstrate that studies to the contrary may actually be supportive of the traditional view when such misinterpretations are dispelled.

The common misunderstandings identified by Ingram include first, that complete comprehension precedes production and second, the gap between comprehension and production is systematically long and predictable. These hypotheses are characterized as overstatements of the actual position defined by Ingram (1974) as "some comprehension of a specific grammatic form or construction occurs before it is produced" (p. 316). Factors that are cited by various authors as evidence contrary to this view are identified and criticized by Ingram. These include (1) the occurrence of overgeneralizations of words and grammatic forms (2) the discrepancy between order of appearance of grammatical

forms and constructions in comprehension and production (3) the observation that in some cases the comprehension of syntactic form is the same as production (4) the use of forms with no apparent understanding and (5) the results of experimental studies that show production to be in advance of comprehension. Factors (3) and (5) are especially relevant to this discussion as they challenge the traditional hypothesis that syntactic comprehension precedes production.

First, in support of his position, Ingram cites studies by Ingram (1971) and Greenfield, Smith and Laufer (1973) that postulate semantic intentions of a relational nature during the holophrastic period. While neither study attributes specifically syntactic knowledge to the holophrastic child, Ingram (1974) believes that "the postulation of these semantic relations in children's grammar implies that there is an understanding of them in adult speech" (p. 319). For children of the telegraphic stage, evidence is cited from Shipley et al. (1969) to support the notion that the child understands more than she can produce. This he believes is demonstrated by the telegraphic speaker's preference for adult form commands as determined by the Shipley et al. (1969) experiment.

These studies and their conclusions have been criticized by Bloom (1974) and in previous sections of this paper. For example, citing the proposals of Greenfield et al. (1973) as proof in favor of his hypothesis would

appear premature since the existence of relational meaning prior to syntax is an issue far from resolution. In effect, Ingram has drawn as evidence proposals that are still open to question and validation. However, Ingram insists that such evidence as Bloom (1974) and others provide does not contradict the traditional view as defined. In sum, arguments that comprehension and production may be closer together than originally supposed for certain constructions does not deny the precedence of comprehension. We are reminded that the position supported by Ingram makes no claim on the length and predictability of the gap between the two processes which is the issue believed to be at the root of Bloom's dissension.

A similar defence is posed by Ingram in response to Fernald's (1972) criticism of the Fraser et al. (1963) study which provides evidence in support of the traditional view. According to Ingram, for the traditional position to be contradicted, one would need evidence to show production consistently in advance of comprehension. One study purporting to demonstrate such a relationship (Keeny and Wolfe, 1972) is criticized by Ingram on the grounds that the comprehension task design did not validly assess comprehension ability.

In sum, Ingram (1974) casts doubt on much of the experimental literature that addresses the comprehension-production relation by questioning the validity of comparing per-

formance on contrived comprehension and production tasks when valid measurement and task equality cannot be ensured.

Bloom (1974;1978)

Bloom (1978) contends that the tendency of child language studies to take comprehension for granted has resulted in the assumption that the developmental relationship between comprehension and production is simply that comprehension precedes production at every stage of development. While conceding that understanding and speaking cannot be considered as completely separate achievements, Bloom takes exception to the view that each represents merely a different mode of the same fundamental development. Instead, Bloom (1978) proposes that

> "the two represent mutually dependent but different underlying processes, with a resulting shifting of influence between them in the course of language development... the developmental gap between comprehension and production probably varies among different children and at different times and may often be more apparent than real" (p. 238).

Evidence in support of her view is provided from diary accounts of acquisition and several experimental studies aimed at comparing comprehension and production. In sum, for both the single word stage and that of multiword utterances, no simple one to one relationship between the two processes is clear. For example, the first words understood by the child are not necessarily the first produced (Leopold, 1939). Similarly, it is not the case that production depends on prior comprehension for each instance in which a word is used. The existence of overextension of

reference for words used by the child hints at the possibility that she may learn to understand the word by learning how to use it ie. generalizing or associating its properties to new situations (Bloom, 1973; Clark, 1973).

In establishing her position regarding the issue of whether the child who is producing only single words processes syntactic information in comprehending speech, Bloom relies on evidence that demonstrates the multidetermined nature of comprehension. According to Bloom, that children during the first two years appear to understand a great deal more than they say does not in itself provide evidence that syntactic processing is involved. In fact, she proposes that syntactic processing is least significant. Because the speech addressed to children at this age is generally confined to comments about the here and now, the child's apparent understanding results from the redundancy of the linguistic message with the supporting context eg. stress, gesture, tone, repetition, exaggeration. She further identifies an asymmetry between the later understanding and speaking of multiword utterances which is based on the observation that

> "children do not have to process syntax to understand reference to relations in immediate events, but children do need to learn something about syntax of the language in order to talk about such relations in any coherent way... thus, knowing a word and knowing a grammar, and understanding structured speech and using structured speech, apparently represent different mental capacities ... it may be misleading to consider that such capacities develop in linear temporal relation, with comprehension simply preceding production (p. 244).

From her review of several experimental studies addressing the comprehension-production question, Bloom concludes that the results obtained are open to multiple interpretation. However, some generalizations appear supportable.

Both understanding and producing a linguistic form (word, grammatic morpheme, syntactic structure) depends on multi-linguistic and nonlinguistic variables. Events in the communicative context interact with information in memory to represent the meaning of messages for the child. The interaction of such components appears different for the processes of comprehension and production and may vary according to the experience of the child and her developing linguistic and cognitive capacities (Clark, 1973).

What a child understands of a particular form may heavily depend on the context in which it is heard. Similarly, the child's use of a form cannot be taken as absolute evidence the child understands the form to the extent that she is able to use it in unlimited reference and in any context.

Comprehension and production tasks created in experimental studies appear to tap different aspects of the child's knowledge of language. Children demonstrate different strategies in performing comprehension and production tasks. The strategies for each appear to vary as a function of the child's linguistic and cognitive development and prior experience (Huttenlocker, 1968a;1968b;1971; and Chapman, 1977).

Summary

It is difficult to assess both the experimental studies and the theoretical positions presented in the previous sections and arrive at a resolution of the comprehension-production issue. The experimental studies have produced contradictory findings as a result of methodological variation and in many cases, weakness. In turn, the theoretical stances of Bloom and Ingram rely on the experimental literature for credibility. For example, Bloom (1978) cites Chapman and Miller (1975) as evidence contrary to Ingram's (1974) position. The next section will provide a detailed, critical review of this study to demonstrate the methodological and interpretive pitfalls to which investigations addressing this issue are prone.

A Detailed Review: Chapman and Miller (1975)

Chapman and Miller (1975) challenge the traditional view set forth by Ingram (1974) that comprehension of syntactic structure precedes production. In fact, it is their prediction that for the young child, syntactic production may actually precede comprehension based on syntactic form alone. Specifically proposed is that the appropriate ordering of subject-verb-object will appear in the child's utterances earlier than the time when word order is used exclusively by the child as a comprehension strategy. In theory, these authors find close allies in Bloom (1978) and Clark, Hutcheson and Van Buren (1974) who also question

the existence of a strict developmental ordering for comprehension and production of syntactic form.

Subjects and Groups: Fifteen children drawn from middle to upper-middle class families were divided on the basis of MLU (computed as per Brown, 1973) into three groups. Table 2 summarizes the age, sex and MLU for all subjects and their group assignment.

Test Sentences and Stimulus Objects: The 24 sentences used in the comprehension and production tasks are listed in Table 3. They were constructed from three animate nouns, three inanimate nouns and six transitive verbs. Care was taken to ensure that the objects and actions required were those familiar to the young child. Stimulus objects included a wooden car, dump truck, sailboat, flexible boy and girl dolls, and a plastic dog. The animate objects were smaller in size and more colorful than the inanimate objects.

Comprehension and Production Tasks: For the comprehension task, the six toys were arranged in front of the subjects. A test sentence was presented preceded by the verbal instruction "Do what I say, make the ..." and the child was required to pick out the two relevant toys and demonstrate the relation encoded in the sentence. In the production task, the child observed the experimenter perform an action with two of the toys and was asked to describe the action performed via these directions: "What's happening... what's going on... what am I doing?". Pretesting ensured the subjects knew the task requirements.

Table 2. Chapman and Miller (1975): Data Summary

					COMPREHENSION TASK			PRODUCTION TASK		
	S_s	MLU	AGE	SEX	SCORABLE RESPONSES	CORRECT RESPONSES	(%)*	SCORABLE RESPONSES	CORRECT RESPONSES	(%)
Group I										
	C	1.89	2;1	M	24	13	(54.2)	16	13	(81.2)
	K	2.01	2;2	F	18	13	(72.2)	11	8	(72.7)
	H	1.70	2;1	F	17	11	(64.7)	17	16	(94.1)
	R	1.53	1.8	M	15	12	(80.0)	7	7	(100.0)
	L	1.76	2;2	F	18	11	(61.1)	13	12	(92.3)
Mean		1.78	2;0		18.4	12	(66.4)	12.8	11	(88.1)
Group II										
	A	2.56	1;11	F	8	5	(62.5)	3	3	(100.0)
	J	2.33	2;1	F	14	11	(78.6)	3	3	(100.0)
	K	2.20	1;11	M	21	13	(61.9)	14	8	(57.0)
	M	2.39	1;8	F	21	13	(61.9)	10	10	(100.0)
	R	2.40	2;0	F	20	15	(75.0)	10	8	(80.0)
Mean		2.38	1;11		16.8	11.4	(68.0)	8	7	(87.4)
Group III										
	B	2.92	2;5	F	19	14	(73.7)	6	5	(83.3)
	S	2.84	2;7	F	21	12	(57.1)	20	17	(85.0)
	C	2.84	2;4	F	20	13	(65.0)	11	9	(81.8)
	K	2.97	2;4	F	23	20	(87.0)	13	13	(100.0)
	A	3.11	2;8	F	19	15	(78.9)	13	11	(84.6)
Mean		2.94	2;6		20	15	(72.3)	13	11	(86.9)

* correct responses divided by scorable responses

Table 3: Chapman and Miller (1975): Stimulus Sentences

A Sentences	B Sentences
The boy is hitting the girl	The girl is hitting the boy
The girl is carrying the dog	The dog is carrying the girl
The dog is chasing the boy	The boy is chasing the dog
The dog is chasing the car	The car is chasing the dog
The boy is carrying the truck	The truck is carrying the boy
The girl is pulling the boat	The boat is pulling the girl
The boat is hitting the girl	The girl is hitting the boat
The truck is bumping the dog	The dog is bumping the truck
The car is pushing the boy	The boy is pushing the car
The truck is pulling the boat	The boat is pulling the truck
The boat is bumping the car	The car is bumping the boat
The car is pushing the truck	The truck is pushing the car

Procedures and Scoring: Two testing sessions and an optional third were conducted with each subject. In the first session, twelve sentences (either A or B set) were presented in the comprehension task in random order for each child. For the production task, the twelve remaining sentences were demonstrated in random order for each child. The second session was identical to the first except the sentences for each task were exchanged. The optional third session consisted of retesting items deemed unscorable due to inattention, unresponsiveness or uncooperativeness. Responses to the two tasks were scored for preservation of the subject-object relation specifically. A summary of responses categories and scoring criteria for each task appears in Table 4. Scorable responses included categories (1) and (2). Those responses in categories (3) and (4) were retested in the third session and the child's response to task accepted as data. The guessing rates for each task are said to be 0.50.

Analysis and Results: The data obtained appears in Table 2. For each task, the child's percent correct score was calculated by dividing the total number of correct responses by the total scorable (right plus wrong) responses.

For each group, a one tailed correlated t-test, degrees of freedom of four, was run on the percent correct scores. The difference between the production and comprehension scores was significant for every group with t-test values equalling 3.77 ($p \leq 0.01$), 2.26 ($p \leq 0.05$) and 3.83 ($p \leq 0.01$)

Table 4. Chapman and Miller (1975): Scoring Criteria

Response Category	Comprehension	Production
(1) Correct	Action appropriately demonstrated with appropriate subject-object assignments	Sentences of these forms: subject-verb-object; subject-verb; verb-object; subject-object
(2) Wrong	Action appropriately demonstrated with subject-object assignment reversed	Sentences of these forms: object-verb-subject; verb-subject; object-verb or object-subject
(3) No Response	Child does not respond	Child does not respond
(4) Undecidable	Child's attention to word order unassessable from response	Verb only; subject only; object only, irrelevant response

respectively for Group I, II, and III. In addition, those instances in which a child gave a scorable response to a sentence on both the comprehension and production tasks were tallied according to the correctness of each response. This data appears in Table 5.

A further analysis was conducted on the number of correct responses to each sentence type for both tasks. Percent correct scores were tallied and appear in Table 6, for comprehension and Table 7, for production. For the comprehension task, performance is consistently near 100 percent for the sentence type possessing an animate subject and inanimate object. For the shorter MLU groups, performance on the sentence type having an inanimate subject and animate object is below 50 percent, indicating that the children were demonstrating the reverse order for the majority of these sentences. Performance on the remaining sentences types is intermediate. For the production task, little variation in percent correct scores occurs among the sentence types. Scores are high across all sentence types and are thus not patterned in the same fashion as the comprehension scores, except at ceiling levels.

Conclusions: Chapman and Miller (1975) conclude from their data that

> "the pattern of performance for each group is clear: correct word order for subject and object is preserved significantly more often in speaking than in serving as a cue to subject and object in a comprehension taks... to couch this finding in the language of other studies, production precedes comprehension in grammatic acquisition for subject-object structure" (p. 362).

Table 5. Chapman and Miller (1975): Percent Correct for Sentences Scorable on both Tasks

Group	Comprehension	Production
I	40 (22/55)	60 (33/55)
II	76 (28/37)	81 (30/37)
III	70 (42/60)	88 (53/60)

Table 6. Chapman and Miller (1975): Comprehension of Sentence Types: Mean Percent Correct by Group.

Subject-Object	Group I	Group II	Group III	$\bar{X}$
Animate-Animate	76.2	71.4	52.0	66.5
Inanimate-Inanimate	53.8	71.4	70.3	65.2
Animate-Inanimate	95.7	90.0	95.8	93.8
Inanimate-Animate	36.4	40.9	73.0	50.1

Table 7. Chapman and Miller (1975): Production of Sentence Types: Mean Percent Correct by Group.

Subject-Object	Group I	Group II	Group III	$\bar{X}$
Animate-Animate	80.0	88.9	82.3	83.7
Inanimate-Inanimate	90.5	75.0	81.8	82.4
Animate-Inanimate	89.5	81.8	87.5	86.3
Inanimate-Animate	85.7	87.5	94.7	89.3

In sum, these authors contend that the English speaking child's use of word order information as a cue to subject-object status is limited and acquired late in contrast to her ability to encode this relation in speech. Further, the authors believe that the examination of patterns of performance for sentence types in comprehension and production shows that the strategies for comprehending, lacking context are different from the strategies for producing sentences, given context. Chapman and Miller (1975) propose that

> "less linguistically advanced children decode subject-object relations in sentences lacking referential support on the basis of a lexical-semantic strategy; the animate noun is assigned subject status and the inanimate noun, object status" (p. 367).

Critical Analysis of Chapman and Miller (1975)

The following is a summary of the major methodological and analytic flaws apparent in the Chapman and Miller (1975) experiment.

First, for all groups, the mean number of scorable responses for the comprehension versus production tasks was widely disparate. Due to this, a comparison of percent correct scores is invalid as the metric for each is different. To clarify, consider the mean difference between the number of scorable responses for the two tasks which appear in Table 8. An examination of the data found in Table 2, shows that each subject gained more scorable responses on the comprehension versus production task, with the exception

Table 8. Chapman and Miller (1975): Mean Difference in Number of Scorable Responses Between Tasks

Group	Comprehension	Production	Difference
I	18.4	12.8	5.6
II	16.8	8.0	8.8
III	20.0	13.0	7.0

of Subject H of Group I. Percent correct scores were computed by dividing the number of correct responses by the number of scorable responses for each task. Since the number of scorable responses for each task was different, a comparison of percent correct scores represents an invalid treatment of the data. The following will clarify this point. Consider the results of Group II found in Table 9.

Table 9. Chapman and Miller (1975): Group II: Scorable Versus Correct Responses

	Comprehension			Production		
Subject	Scorable Responses	Correct Responses	%	Scorable Responses	Correct Responses	%
A	8	5	62.5	3	3	100.0
J	14	11	78.6	3	3	100.0
K	21	13	61.9	14	8	57.0
M	21	13	61.9	10	10	100.0
R	20	15	75.0	10	8	80.0
Mean	16.8	11.4	68.0	8	7	87.4

A superficial examination of the data via comparison of percent correct, indicates as Chapman and Miller conclude that performance on the production task far surpassed that of the comprehension task (87.4% versus 68.0%, respectively). However, to compare percent correct scores when the number of scorable responses (ie. the metric) is not equal, is a spurious procedure especially when N values are small. Take for example, Subjects A, J and M. These subjects were scorable on only three, three and ten, respectively, of the 24 production trials. For those trials deemed scorable,

100 percent accuracy was demonstrated. However, the actual number of correct responses in each case is less than the actual number of correct responses on the comprehension task which equalled five, eleven and thirteen respectively. The lower percentage scores on the comprehension task result because the number of scorable responses was greater than on the production task. Since Group II was comprised of only five subjects, any single subject could greatly influence the mean percent correct score for the group by either excelling or failing. In this case, Subjects A and J greatly enhanced the mean percent correct for Group II on the production task by happening to score with 100 percent accuracy on the mere three trials (out of a possible 24) considered to be scorable. That they did so may be purely attributable to chance.

Further, the use of percent correct may in itself be misleading when calculated on small values. In fact, small real differences may be exaggerated by using percent correct scores. This can be easily demonstrated by the following example. Consider the scores for the production task for Group II found in Table 9. If each score is reduced by merely one point, the mean percent correct drops to 67. If each score is reduced by two points, the mean percent deflates to 50. If reduced by three points, the mean percent correct literally dissolves to 31. Thus, a real difference of only three points produces a difference of 56 percent. One can easily see how the use of percents

in this case can lead to an exaggeration of actual differences.

In sum, a precondition to comparing percent scores for two tasks is that the number of trials (in this case, scorable responses) be equal. This condition was clearly not met by Chapman and Miller (1975) as only two of 15 subjects had equal or nearly equal scorable responses for each task. Further, the use of percent correct may in itself be a misleading procedure since it tends to exaggerate small actual differences in scores.

Secondly, the subset of data in which the number of scorable responses to the test sentences was the same for both tasks does not exhibit a significant difference in favor of production for all groups. Consider the data presented in Table 5, where the number of scorable responses to both tasks equalled 55, 37, and 60, respectively for Group I, II and III. Group II shows no appreciable difference in performance between the two tasks (30 of 37 correct for production versus 28 of 37 correct for comprehension). In addition, the authors did not perform a statistical test of significance on this data, preferring instead to identify "patterns" or trends in performance in favor of production.

Thirdly, when the Fernald (1972) correction is applied to their data, no significant difference is found between performance on the two tasks. To recall, Fernald (1972) and Baird (1972) criticized the number correct scoring procedure used in the ICP studies. It was demonstrated that

the probability of correctly guessing was biased in favor of the comprehension task. These authors proposed that studies comparing comprehension and production must employ scoring procedures characterized by (1) equivalent guessing rates for both tasks and, (2) scoring criteria specific to the syntactic form in question. The scoring procedures used in the Chapman and Miller (1975) experiment meet these criteria only if one assumes, as they state, that "the exclusion of unscorable responses does not differentially bias the probability estimate obtained from the scorable responses" (p. 363). This point warrants clarification. The authors assume that the exclusion of unscorable responses does not affect the probability rates for guessing on the scorable items. In effect they assume that "unscorable responses do not signal an internal knowledge state in the organism any different from that signalled by the scorable responses" (p. 364). They contrast this with what they believe to be the assumption underlying the Fernald (1972) correction, specifically, that the random assignment of 50 percent correct - 50 percent incorrect designations to unscorable data is "tantamount to assuming that the subject is guessing, and therefore, not able to attend to order on any of the unscorable items" (p. 363). Since the number of unscorable responses was high and because the children earned several unscorable responses simply through inattentiveness, Chapman and Miller believe the application of Fernald's correction to be inappropriate and distortive of

the data. This conclusion may be questionable if the following facts are examined.

Consider Subject A of Group II who produced scorable responses for only three of 24 production trials (see Table 9). Twenty-one of 24 trials are thus excluded from the data. In total, for all subjects, 193 of the 360 production trials were unscorable and hence not considered in the analysis. This is in contrast to 82 of the 360 trials deleted from the comprehension task. If the task requirements were indeed equal and if unscorable responses do not signal an internal knowledge state any different from that signalled by scorable responses as Chapman and Miller assume, what explanation can the authors provide to account for the considerable difference between the number of scorable responses for the two tasks? Further, how can one justify proceeding with an analysis in which over half of the trials for one task have been deleted?

Finally, that significantly fewer scorable responses were realized for the production task may indicate that the task requirements for each task were unequal. The authors fail to offer any explanation for the relatively poor rate of scorable responding to the production trials. If the task requirements were equal, one would expect the number of scorable responses for each to be comparable. This is not the case as the subjects were scorable on significantly more of the comprehension trials (278 versus 167).

In view of what can be considered serious analytic and interpretive problems inherent in the Chapman and Miller (1975) design, a replication of their findings must precede an acceptance of their conclusions. What follows is such an attempt.

CHAPTER THREE

HYPOTHESES AND DEFINITIONS

The present study is an attempt to test the predictions made by Chapman and Miller (1975) that (1) the child's productive control of the subject-verb-object relation will exceed comprehension performance based on syntactic form alone, and (2) less linguistically advanced children decode subject-object relations lacking referential support on the basis of a lexical-semantic strategy which assigns animate nouns to subject status and inanimate nouns to object status. Four hypotheses have been formulated to test these predictions.

Hypotheses

Hypothesis I: There will be no difference in error scores between Group I and Group II on tasks or sentence types.

Hypothesis II: There will be no difference in error scores between the comprehension and production tasks for groups or sentence types.

Hypothesis III: There will be no difference in error scores among sentence types for groups or tasks.

Hypothesis IV: There will be no difference in the number of scorable responses between groups or tasks.

A three-way analysis of variance will be used in testing Hypotheses I, II and III. For Hypothesis IV, a two-way analysis of variance will be employed.

Definitions

Groups: The ten subjects were divided on the basis of MLU into two groups of five. The mean MLU and age as well as the age and MLU range for each group is found in Table 11.

Tasks: Each subject was required to perform 48 trials for each of the comprehension and production tasks. The comprehension task required the child to demonstrate the relation between a given subject and object encoded in the test sentence. The production task required the child to verbally state the subject-object relation demonstrated by the experimenter. A full description of task requirements is found in the following chapter.

Sentence Types: Four sentence types were constructed on the basis of the animacy-inanimacy of the subject and object. These appear in Table 10. Twelve of each sentence type were tested for both the comprehension and production tasks.

Table 10. Sentence types

Sentence Type	Subject	Object	Example
1	Animate	Animate	Boy hit girl.
2	Inanimate	Inanimate	Truck hit car.
3	Animate	Inanimate	Boy hit car.
4	Inanimate	Animate	Car hit boy.

Error Scores: An error score represents the total number of responses that fall into the wrong categories of Table 4 for comprehension and production.

Scorable Responses: Responses defined as scorable fall into the right and wrong categories of Table 4 for comprehension and production. It should be noted that only scorable responses are considered in the analysis.

CHAPTER FOUR

EXPERIMENTAL PROCEDURE AND DESIGN

The Chapman and Miller (1975) experiment has been shown to possess serious methodological and analytic deficiencies. The present study has introduced changes in procedures and analysis in an attempt to rectify such deficiencies. The major modifications are summarized below.

(1) MLU calculation

The number of utterances on which MLU values are based is constant for each subject (100). In addition, further spontaneous utterances were collected to provide additional naturalistic production for comparison with production test performance. In the Chapman and Miller (1975) experiment, the number of utterances included in the calculation of MLU ranged from 50 to 150 and was thus not constant. Further, although adherence to Brown's (1973) method of MLU computation is claimed, Brown stipulates at least 100 utterances must be included for calculation. This is not demonstrated in the Chapman and Miller (1975) experiment. Thus, MLU values may not reflect accurately some subjects' productive level in the Chapman and Miller (1975) experiment.

(2) Comprehension testing

The procedures for comprehension testing have been simplified by reducing the number of toys available to the child for any one test sentence from six to two. In the comprehension testing of the Chapman and Miller (1975) experiment,

six toys were placed in front of the child and the test sentence provided. The child was then required to select the two appropriate toys and demonstrate the relation stipulated. If the distractibility of young subjects is considered, having six toys available for manipulation may provide a diversion that interferes with comprehension performance. For this reason, only the two relevant toys were in sight during comprehension testing in the present study.

(3) Number of trials

The number of trials for each task has been increased from 24 to 48. Increasing the number of trials for each subject allows for a greater sampling of her performance and reduces the role of chance variation in producing significant results. Also, by increasing the number of trials to 48, each noun is used an equal number of times as either the subject or object.

(4) Verbs used

Only three of the original six verbs used in the Chapman and Miller (1975) experiment have been used in the present study. The verbs chase, carry and pull were thought to involve physical actions more difficult than push, hit and bump. Since the relation of subject and object is the variable under study, the simplest actions required by the verb form were considered the least confounding. For the same reason, the verb form was verbally provided to the child during the demonstrations of the production trials.

(5) Data analysis

The analysis of data obtained is based upon a comparison of raw scores rather than percents. The previous section has outlined the obvious danger in using percent scores in an analysis for which the metric is not constant. In addition, it was demonstrated that percent scores may in fact exaggerate small real differences. For these reasons, only raw scores were considered in the analysis.

(6) Statistical tests used

The statistical tests used in the present analysis are first, a three-way analysis of variance with two repeated measures for Hypotheses I, II and III, and secondly, a two-way analysis of variance for Hypothesis IV. This contrasts with the use of repeated, one-tailed t-tests found in Chapman and Miller (1975). The use of individual t-tests for each group increases the likelihood of committing a Type I error; that is, such procedures increase the level of significance and thus the chance of finding significant differences. The F statistic is a more conservative treatment which maintains the alpha level at a constant value.

Subjects and Groups

Ten children drawn from middle class homes participated. The subjects were grouped on the basis of mean length of utterance (MLU) into two groups of five. Group I children correspond to Stage II development as per Brown's (1973) classification with Group II approximating Stage IV. The

age, sex and MLU in morphemes for each subject appear in Table 11.

Table 11. Sex, Age and MLU Values for all Subjects

Group	Subject	Sex	Age	MLU[1]	(MLU)[2]	Total Utterances
I	A	M	2;10	2.14	(2.87)	257
	B	F	2;4	2.19	(2.49)	274
	C	M	2;2	2.41	(2.38)	253
	D	F	2;3	2.69	(2.79)	248
	E	M	2;5	2.80	(2.88)	233
Mean			2;5	2.45	(2.68)	253
II	F	F	2;8	3.23	(3.05)	197
	G	F	2;10	3.56	(3.66)	170
	H	F	2;7	3.73	(3.93)	180
	I	F	2;7	4.06	(3.86)	203
	J	M	2;10	4.17	(3.98)	214
Mean			2;8	3.75	(3.70)	193

[1]MLU: based on 100 utterances

[2](MLU): based on total utterances obtained

Speech Sample

Computation of MLU was based on a tape recorded spontaneous speech sample obtained during the experimenter's visit to the child's home. Verbal interaction between the child, mother, siblings and experimenter yielded a range of 170 to 274 utterances for computation. Computation was conducted using Brown's (1973) procedures without modification. For some subjects, additional spontaneous sampling

was conducted in order to compare the child's production test performance with naturalistic production. The MLU values in Table 11 are based on the first one hundred utterances occurring after the first page of transcription. Bracketed figures in the MLU column represent the MLU value based on the total sample obtained and were not used in determining group assignment. The MLU values based on the first one hundred utterances correspond closely to the values calculated on the entire sample obtained and thus reliably estimate the subjects' productive level.

Sentences

Forty-eight sentences were constructed using three familiar animate nouns (boy, girl, dog), three inanimate nouns (truck, car, boat) and three transitive verbs (push, hit, bump). The verbs allow both animate and inanimate subject and object. The sentences used appear in Table 12. All sentences are active, reversible and in the present tense. The same sentences are used in both the comprehension and production tasks. Each sentence type corresponds to the defining characteristics outlined in Table 10.

Table 12: Test Sentences for the Comprehension and Production Tasks

	TYPE 1	TYPE 2	TYPE 3	TYPE 4
1.	Girl bump dog	Car bump boat	Girl bump car	Car bump girl
2.	Girl bump boy	Car bump truck	Boy bump car	Car bump boy
3.	Boy hit dog	Boat push car	Dog push boat	Boat push dog
4.	Boy hit girl	Boat push truck	Boy push boat	Boat push boy
5.	Dog push girl	Truck hit car	Girl hit truck	Truck hit girl
6.	Dog push boy	Truck hit boat	Dog hit truck	Truck hit dog
7.	Dog bump girl	Boat bump car	Dog bump car	Car bump dog
8.	Boy bump girl	Truck bump car	Girl push boat	Boat push girl
9.	Dog hit boy	Car push boat	Boy hit truck	Truck hit boy
10.	Girl hit boy	Truck push boat	Boy push car	Car push boy
11.	Girl push dog	Boat hit truck	Dog bump truck	Truck bump dog
12.	Boy push dog	Car hit truck	Girl hit boat	Boat hit girl

Stimulus Objects

The six toys used in both tasks were items produced by Fisher Price Toys Limited. Included for the animate nouns were boy, girl and dog,"Little People" of uniform size and composition. The inanimate noun toys were a plastic car, truck and boat of closely similar size. The animate noun toys and the inanimate noun toys were proportionate in size, with the inanimates larger in bulk than the animates.

Comprehension Task

The child was positioned in front of and facing the tester. Two toys were placed in front of the child at equal distance from each hand. Attentional cues, including verbal ("Watch and Listen") and nonverbal (touching) stimuli preceded the presentation of each test sentence. The child's task was to demonstrate the sentence provided. Verbal instructions given to the child included "Do what I say, make the girl hit the boy" or "You do it, make the car hit the boat" and were repeated until the child responded. Care was taken to deliver the test sentences without intonation or gesture. If no response was given, the attentional cue and verbal instruction were repeated. Preceding each testing session, three practice items were presented with feedback provided to the child informing her she was "playing the game" correctly. Whether the child demonstrated the sentence with the correct or incorrect subject-object assignment was not made known to the child either in

practice or testing sessions. Reinforcement regarding performance was related only to sitting and attending and was restricted to general comments such as "You're playing my game nicely" or "You're watching and listening- good girl". Because only two toys were used for each task, the positioning of the subject toy was randomly varied either to the child's left or right for each demonstration.

Production Task

The child sat facing the tester while the tester performed an action (pushing, hitting, bumping) with two of the toys. Each demonstration was preceded by attentional cues including "Watch what happened- I want you to tell me what happens" or "Look (demonstration of action) - what happened, you tell me what happened". Each demonstration of the test sentence was accompanied by verbal statement of the verb being performed. Reinforcement during testing was confined to comments similar to those used in the comprehension sessions. At no time was the child reinforced for correctly encoding the subject-object relation.

Pretesting

During the session conducted to obtain each child's spontaneous speech sample, pretesting to ensure the child's knowledge of the lexical items used in the test sentences was performed. The child was first required to point to each toy as it was named and then to name each toy as the tester pointed to it. Each verb was verbally labelled and the

child then required to demonstrate its action. The action of the verb was performed by the tester and the child was required to identify the action. Three subjects demonstrated confusion between boy and girl so "big bird" was substituted for either boy or girl where indicated. All subjects demonstrated consistent familiarity with all verbs and nouns used.

Scoring

The scoring procedure found in the present study is a duplicate of that used in Chapman and Miller (1975).

Comprehension Scoring

Responses to the comprehension task were scored for the comprehension of the subject-object relation specifically. Response categories included correct, if the action was demonstrated with the correct subject-object assignment; wrong, if the subject-object assignment was reversed; no response; and undecidable, if the child's attention to word order could not be assessed from the response eg. if the child placed the boy in the truck or if she simultaneously banged the two toys together. Chance success for the comprehension task if only scorable responses are considered as data is 0.50. A summary of the comprehension scoring criteria appears in Table 4.

Production Scoring

Responses to the production task were scored specifically for the appropriate ordering of subject and object with respect to the verb. Response categories included correct, including subject-verb-object, subject-verb, verb-object and subject-object sentences; wrong, including object-verb-subject, verb-subject, object-verb and object-subject sentences; no response; and, undecidable, including verb only, noun only, and all other sentences from which attention to the ordering of subject and object could not be determined. The guessing rate for correctly ordering subject and object in scorable responses is 0.50. Table 4 presents a summary of the scoring criteria for production used in the present study.

Procedures

Five sessions were conducted with each subject. The first session consisted of obtaining the spontaneous speech sample used for the MLU computation and for pretesting the child's familiarity with lexical items found in the test sentences. Testing was conducted during the remaining four sessions. All testing sessions were completed within a maximum of 27 days from the first MLU sampling session.

Each testing session was comprised of twelve comprehension and twelve production sentences. The ordering of the comprehension and production tests was alternated for each session. Each set of twelve sentences contained three instances of each sentence type, with the order of

presentation randomly determined for each subject and session. If during the testing session the child became uncooperative, the session was terminated and resumed on a subsequent visit. Mothers were present for all sessions and were advised not to participate verbally during testing sessions.

Reliability checks for both comprehension and production scoring were conducted with four subjects for one session each. Mothers acted as reliability scorers for the comprehension task while the production portion of the session was tape recorded and validated by an independent listener.

During the comprehension testing, note was made of hand preference, response revisions and questions asked by the child. The child's verbatim responses during production testing were transcribed for all subjects.

Reliability Measures

Reliability checks were run on MLU calculation, comprehension and production scoring.

MLU Calculation Reliability: An independent listener transcribed the entire speech sample obtained from Subject D of Group I and calculated MLU using Brown's (1973) procedures.

The number of utterances and morphemes counted by the experimenter equalled 170 and 624 respectively, yielding an MLU value of 3.67. This is compared with the independent listener's calculation of 189 utterances totalling 637

morphemes yielding an MLU of 3.37. Reliability for number of utterances counted equals 90 percent; for number of morphemes counted, 98 percent, for MLU overall, 92 percent.

Comprehension Scoring Reliability: For two subjects from each group, a reliability check for comprehension scoring was conducted using the child's mother as an independent scorer. The scores assigned by the experimenter and mother were compared for the twelve sentences comprising one comprehension session. For each of the subjects, no variability in scores assigned was found. Thus, reliability for comprehension scoring was 100%.

Production Scoring Reliability: For two subjects from each group, a reliability check for production was conducted. The child's production responses for the twelve sentences comprising one session were tape recorded and validated by an independent listener. Forty-eight sentences in total were compared for assigned scores. Disagreement in assigned scores was found for 4 items. Thus reliability for production scoring equals 92 percent.

CHAPTER FIVE

RESULTS

Table 13 provides the number of errors scored by each subject on each task and sentence type. In testing Hypotheses I, II and III, a three-way analysis of variance with two repeated measures was run on raw error scores. Table 14 summarizes the main and interaction effects for all factors.

Main Factors

Groups: Overall, no significant difference ($p = 0.119$) in error scores was found between Group I and II.

Tasks: Overall, a significant difference ($p = 0.002$) in error scores was found between the comprehension and production tasks.

Sentence Types: Overall, a significant difference ($p = 0.003$) in error scores was found among sentences (1), (2), (3) and (4).

Interactions

Group by Task: No significant interaction ($p = 0.548$) was found between groups and tasks.

Group by Sentence Type: No significant interaction ($p = 0.550$) was found between groups and sentence types.

Task by Sentence Type: A significant interaction ($p = 0.001$) was found between tasks and sentence types.

Table 13. Data Summary

Group	Subject	Comprehension Errors by Sentence Type					Production Errors by Sentence Type					Scorable Responses	
		1	2	3	4	Total	1	2	3	4	Total	Comprehension	Production
I	A	6	8	11	0	25	0	0	0	0	0	48	48
	B	5	4	4	4	17	0	1	0	1	2	48	38
	C	6	3	7	6	22	0	1	0	1	2	48	37
	D	4	5	11	2	22	2	2	0	2	6	48	46
	E	5	5	10	1	21	6	3	5	9	23	48	43
	TOTAL	26	25	43	13	107	8	7	5	13	33		
II	F	8	5	12	0	25	1	1	0	2	4	48	47
	G	3	7	8	4	22	1	1	2	0	4	48	47
	H	1	5	1	2	9	2	3	0	0	5	48	48
	I	3	3	4	0	10	0	1	1	0	2	48	45
	J	0	2	3	0	5	0	0	0	0	0	48	48
	TOTAL	45	22	28	6	71	4	6	3	2	15		

Table 14. Three-way Analysis of Variance: Main and Interaction Effects ($\alpha = 0.05$)

Source	Sums of Squares	Df	Mean Squares	F	P
A (Groups)	36.450	1	36.450	3.05	0.119
B (Tasks)	211.250	1	211.250	20.56	0.002*
C (Sentence Types)	51.850	3	17.283	5.98	0.003*
A x B	4.050	1	4.050	0.39	0.548
A x C	6.250	3	2.083	0.72	0.550
B x C	87.850	3	29.283	7.57	0.001**
A x B x C	7.850	3	2.617	0.68	0.575

*$p \leq 0.01$

**$p \leq 0.001$

Group by Task by Sentence Type: No significant interaction ($p = 0.575$) was found among all main factors.

Post-hoc Contrasts

Between Tasks for Given Groups: Table 15 provides the post hoc contrasts between tasks for given groups. For Group I, the number of errors scored on the comprehension task was significantly greater ($p = 0.006$) than on the production task. For Group II, the number of errors scored on the comprehension task was significantly greater ($p = 0.025$) than the production task.

Between Tasks for Given Sentence Types: Table 16 provides the post hoc contrasts between tasks for given sentence types. For sentence types (1), (2) and (3), significantly greater ($p = 0.000$ in all cases) errors were scored on the comprehension task. No significant difference ($p = 0.592$) in error scores was found between tasks for sentence type (4).

Table 15. Post hoc Contrasts Between Tasks for Given Groups

Contrast	Mean Difference	Mean Square	Df1	Df2	F	P
Group I: Comp vs Prod.	3.70	10.27	1	8	13.324	0.006*
Group II: Comp vs Prod.	2.80	10.27	1	8	7.630	0.025**

*$p \leq 0.01$ **$p \leq 0.05$

Table 16. Post hoc Contrast Between Tasks for Given Sentence Types

	Contrast	Mean Difference	Mean Square	Df1	Df2	F	P
Sentence Type	1: Comp vs Prod	2.90	5.47	1	32	15.378	0.000*
	2: Comp vs Prod	3.40	5.47	1	"	21.138	0.000*
	3: Comp vs Prod	6.30	5.47	1	"	72.576	0.000*
	4: Comp vs Prod	0.40	5.47	1	"	0.293	0.592

*$p \leq 0.001$

Among Sentence Types for Given Tasks: Table 17 summarizes the post hoc contrasts among sentence types for given tasks. For the comprehension task, the following contrasts were significant: sentence type (1) versus (3) ($p = 0.001$); (1) versus (4) ($p = 0.01$); (2) versus (3) ($p = 0.005$); (2) versus (4) ($p = 0.001$) and, (3) versus (4) ($p = 0.000$). Errors on sentence types (1), (2) and (3) were significantly greater than sentence type (4). In addition, errors on sentence type (3) were significantly greater than all other other types. All remaining contrasts were not significant at the 0.05 level. For the production task, no contrasts were significant at the 0.05 level. In sum, no significant differences in error scores were found among sentence types for the production task.

In testing Hypothesis IV, a two-way analysis of variance was conducted on the number of scorable responses obtained. Table 18 summarizes the main and interaction effects for all factors.

Groups: Overall, no significant difference ($p = 0.073$) in the number of scorable responses was found between Group I and II.

Table 17. Post hoc Contrasts Among Sentence Types for Given Tasks

Contrast		$\bar{X}$ Diff	$\bar{X}$ Sq	Df1	Df2	F	P
Comprehension: Type	1 vs 2	-0.60	3.38	1	48	0.533	0.469
	1 vs 3	-3.00	3.38	1	48	13.317	0.001**
	1 vs 4	2.20	3.38	1	48	7.162	0.010*
	2 vs 3	-2.40	3.38	1	48	8.523	0.005*
	2 vs 4	2.80	3.38	1	48	11.600	0.001**
	3 vs 4	5.20	3.38	1	48	40.010	0.000**
Production: Type	1 vs 2	-0.10	3.38	1	48	0.015	0.904
	1 vs 3	0.40	3.38	1	48	0.237	0.629
	1 vs 4	-0.30	3.38	1	48	0.133	0.017
	2 vs 3	0.50	3.38	1	48	0.370	0.546
	2 vs 4	-0.20	3.38	1	48	0.059	0.809
	3 vs 4	-0.70	3.38	1	48	0.725	0.399

*$p \leq 0.01$ **$p \leq 0.001$

Table 18. Two-way Analysis of Variance: Main and Interaction Effects

Source	Sums of Squares	Df	$\bar{X}$ Square	F	P
A (Groups)	26.445	1	26.445	4.265	0.073
B (Tasks)	54.434	1	54.434	8.779	0.018*
A x B	26.465	1	26.465	4.268	0.073

*$p \leq .05$

Tasks: Overall, a significant difference ($p = 0.018$) in scorable responses in favor of comprehension was found between the two tasks. Thus, subjects were scorable on a significantly greater number of comprehension trials than production trials.

Group by Task Interaction: No significant interaction ($p = 0.073$) was found between groups and tasks.

CHAPTER SIX

DISCUSSION OF RESULTS

Hypothesis I: There will be no difference in error scores between Group I and II for tasks or sentence types.

The results found in Table 14 indicate no significant difference in error scores ($p = 0.119$) between Group I and II. Further, no significant interaction was found between groups and tasks ($p = 0.548$) or groups and sentence types ($p = 0.550$). Thus, Hypothesis I is not rejected.

Subjects were assigned to groups on the basis of MLU and age. Group I had a mean MLU of 2.45 and mean age of 28.8 months while Group II had a mean MLU of 3.75 and mean age of 32.4 months (see Table 11). That no significant difference in performance was found between the groups may cast doubt on the assumption that the groups differed along the dimensions on which they were formed, that is, MLU and age. To determine whether the mean MLU and age of Group I was significantly less than that of Group II, a one-tailed independent t-test with 8 degrees of freedom was performed on both the MLU values and age in months. The mean MLU and age of Group I was found to be significantly less than that of Group II, with t values of -6.05 ($p \leq .01$) for MLU and -2.34 ($p \leq .05$) for age. Thus, the groups did differ significantly on the dimensions on which they were formed.

According to Brown's (1973) classification, the subjects of Group I and II correspond to Stage II and IV

respectively, of grammatic development. Stage levels are determined by the length and relative grammatic and semantic complexity of utterances. It would seem logical to expect a child whose average constructions are both longer and more complex to out-perform a child whose productive development is substantially less. However, the analysis of error scores shows no significant difference in test performance between the groups both for tasks and sentence types. Thus, overall, differences in productive maturity (MLU) and age did not affect test performance. From this, it is tempting to postulate that productive ability (up to Stage IV) is not related to test performance on either task. However, this assumption would require verification. An alternate explanation is found in reviewing the type of data analysis used in the present study. The F ratio is a relatively conservative statistic. Since the degrees of freedom equalled one, large differences would be required to produce significant results. It is possible that a larger N could have produced a different outcome ie. significant group differences. This, again, requires verification.

Hypothesis II: There will be no difference in error scores between the comprehension and production tasks for groups or sentence types.

The results found in Table 14 indicate a significant difference ($p = 0.002$) in error scores between the compre-

hension and production tasks. Overall, more errors were scored on the comprehension than the production task. In addition, a significant ($p = 0.001$) interaction was found between tasks and sentence types. Therefore, Hypothesis II is rejected.

The conclusion reached by Chapman and Miller (1975) that productive control of the subject-object relation exceeds the child's ability to use word order as a cue to subject and object in a comprehension task, would appear to be supported by the present, more conservative analysis. However, before accepting the Chapman and Miller position, comment must be made on the task requirements for the comprehension and production tests.

Ingram (1974) has questioned the validity of comparing test performance on comprehension and production tasks when the equality of task requirements cannot be ensured. An examination of the task requirements for comprehension and production reveals a factor that operates in favor of production performance. Specifically, the subjects had the benefit of a visual cue during the production trials that may have facilitated the correct encoding of the subject-object relation.

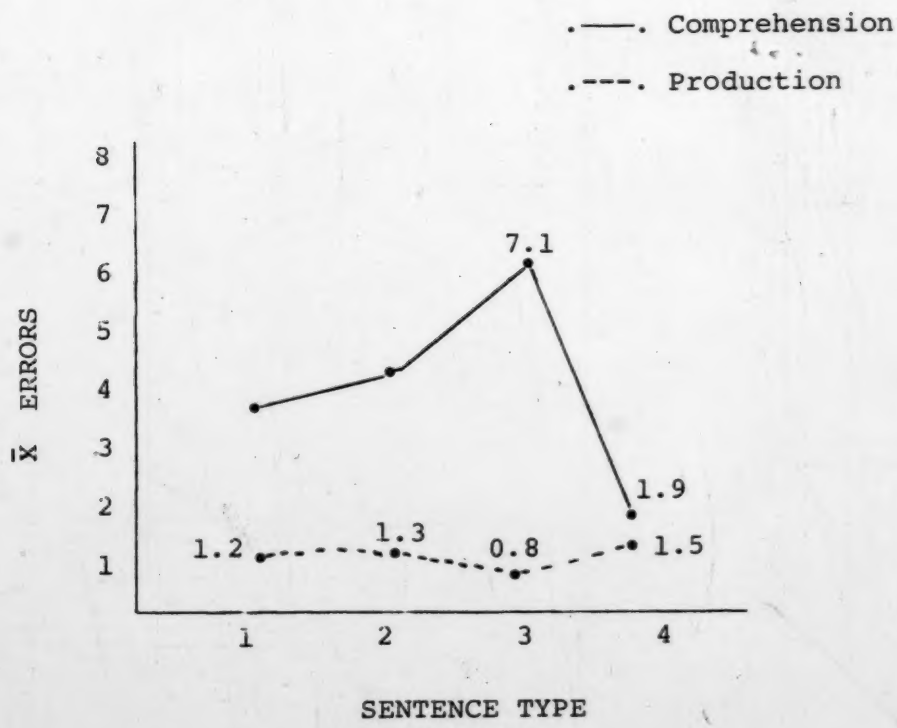
It will be recalled that the production task involved having the tester demonstrate a given subject-verb-object relation after which the child was required to verbally state what she saw occur. The demonstration was such that the subject toy remained in the tester's hand while the

object toy rested on the floor in the position caused by the action performed ie. usually toppled over. It was observed by the experimenter that the subjects would generally use these contextual cues as a reference in assigning subject-object status; that is, the children frequently looked back and forth at the subject toy remaining in the tester's hand and the disturbed object toy on the floor. It was usually after such visual confirmation that the subjects responded. The comprehension task did not allow such visual support; the children could rely only on the verbal instruction issued. The comments of Clark et al. (1974) offer further clarification:

> "If it is true that the situation plays a major part in helping the child to integrate the constituents of utterances then there is reason to expect the child to be at an advantage when he is producing utterances rather than when he is responding to other peoples' utterances, since when he himself is speaking he will be remarking on those aspects of the situation which for him are prominent, whereas when another person is speaking he will have to make the effort to redirect his attention to attain their view of the situation (p. 48-49).

That the additional visual cues inherent in the production testing design introduced a confounding effect may also explain the significant interaction found among tasks and sentence types. Figure 1 graphically illustrates the interaction between tasks and sentence types. As illustrated in Figure 1, the pattern of differences among sentence types for mean errors scored is not the same for the comprehension and production tasks. The post-hoc contrasts among

Figure 1: Interaction Between Tasks and Sentence Types



sentence types for given tasks (see Table 17), indicates that for the comprehension task, errors on sentence types (1), (2) and (3) were significantly greater than (4). In addition, errors on sentence type (3) were significantly greater than types (1), (2) and (4). In contrast, no significant difference in errors scores among sentence types was found for the production task. This interaction may have been the result of unequal task requirements. As stated, the confounding influence of visual cues available in the production task is suspect. Another potentially interfering variable is found within the sentences themselves.

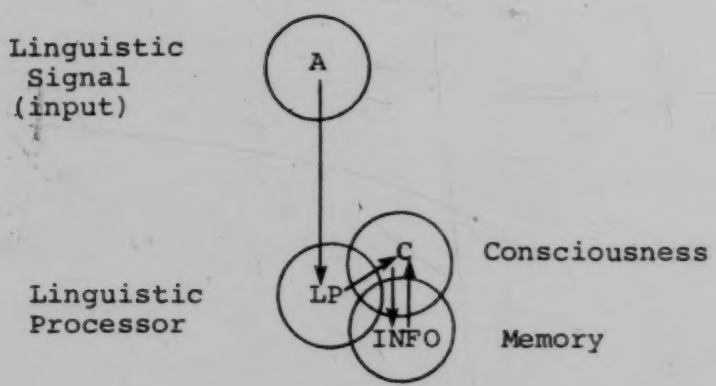
Although syntactically identical, the four sentence types employed differ in semantic likelihood of occurrence; that is, some sentence types reflect likely relations between objects in the real world while others do not. For example, the sentence "car hit boy" (type 4) is more congruent with experiential knowledge than its reverse, "boy hit car" (type 3). Sentence types (1) and (2) could in the sense above be defined as neutral. Evidence can be presented to show that in the comprehension task, subjects were susceptible to a "semantic interference" in which knowledge of what the child knew to be true in the real world interfered with his comprehension performance for those sentence types in which semantically unlikely relations between subject and object were expressed.

It will be recalled that the pattern of differences among sentence types for mean errors scored was not the same for both tasks. In the comprehension task, errors scored on animate-inanimate (type 3) sentences were significantly greater than all other types; conversely, errors scored on inanimate-animate (type 4) sentences were significantly less than the remaining three. Overall, subjects reversed the assignment of subject and object for sentence type (3) on 71 of 120 (59%) of comprehension trials. This contrasts with the performance on sentence type (4) where errors earned equalled only 19 of 120 (16%). Thus, errors scored on those sentences judged to be semantically unlikely (type 3) were far greater than those scored on sentences which expressed relations between objects as they might occur in the child's experience (type 4). This pattern of errors however, is not present in the production task. No significant difference in error scores was found among sentence types.

It would seem then that what the child knew about the relations between objects and events in the real world interfered with her performance on the comprehension but not the production task. The reason for this is likely the fact that during the comprehension task, no contextual support was provided. This contrasts with the production trials in which the physical evidence, that is, the visual cue of an object knocked over, aided the child in assigning subject-object status.

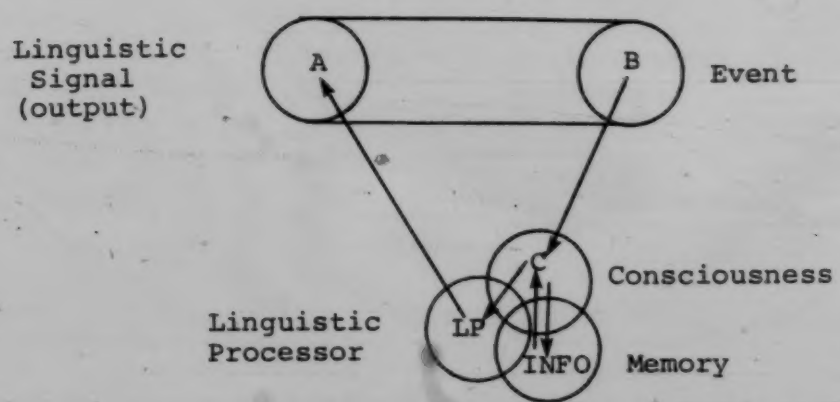
Bloom (1978) has outlined how the processes of comprehension and production may vary with and without contextual support. Figures 2 and 3 respectively schematize the processes of comprehension independent of context and production with the support of context.

Figure 2. Comprehension independent of context



Adapted from Bloom, L. and Lahey, M. (1978) p. 258.

Figure 3. Production with contextual support



Adapted from Bloom, L. and Lahey, M. (1978) p. 255.

Three components interact with each other when the child produces and understands messages: the child's immediate consciousness (C), memory or conceptual knowledge of the world (M) and, linguistic schema, that is lexicon and grammar (LP). These three aspects of the child's thinking are thought by Bloom to interact differently depending on the presence or absence of context.

In the case illustrated in Figure 2, the linguistic signal is the sole external source of information about the message; therefore, information from (LP) is compared with what the child already knows (INFO in MEMORY) in order to obtain a mental representation of the content of the message at (C) and thus understand it. This is contrasted with the situation depicted in Figure 3, where the child is presented with an event (B) and forms a semantic intention based on information from memory and the immediate context which is subsequently linguistically processed to yield an utterance.

Figures 2 and 3 also schematize the comprehension and production tasks used in the present study. It would seem possible that the additional information provided by the context in the production task overrode what the child knew to be the usual relations between the objects in the real world. In contrast, for the comprehension task, the child's experiential knowledge confounded his immediate memory of the relation expressed in the linguistic message alone. Thus, the errors scored on sentence type (3) in the compre-

hension task should be expected to exceed those scored in the production task. This was indeed the case for the present data.

In sum, it would appear that the task requirements for each task were unequal. Thus, the finding that production performance exceeded comprehension may be attributable to the confounding effect of the visual cues available during the production trials.

Before proceeding with the discussion relevant to Hypotnesis III, some comment should be made regarding the unique performance of Subject E of Group I. This child reversed the majority trend by scoring more errors on the production than the comprehension task. In addition, Subject E responded with the complete subject-verb-object structure for 31 of 43 scorable production trials. This contrasts with the other subjects of Group I who produced subject-verb and verb-object structures for all production trials. Of the 31 complete (subject-verb-object) utterances, 20 were produced with the ordering of subject and object reversed. The trials in which partial (verb-object; subject-verb) responses were produced netted only three reversed assignments of subject and object. The same trend was apparent in this child's spontaneous utterances where several instances of word order confusion were noted on three to five word utterances. For example, after having driven his tricycle into a chair he was asked to tell what happened. He replied "rocking chair hit a car!" Other

examples include "I want ball catch", "cry a people", "in there people" and "a people up here sit". However, instances of correct ordering of the subject and object relation were also noted including "a car hit boat".

Since this subject had the highest MLU of Group I (2.80), it would seem possible that he was in the midst of a transition in which the additional knowledge required to correctly combine the constituent structures of subject-verb and verb-object into the composite subject-verb-object structure was not yet complete. This could account for his accurate productive control of the subject-verb and verb-object structures compared to his inconsistent success with appropriately ordering the subject-verb-object relation in both the testing and natural settings. This finding provides empirical support for the "law of cumulative complexity" outlined by Brown and Hanlon (1970). Brown (1973) explains:

> "The general prediction then is that any child who produces some more complex construction will also produce the simpler component constructions into which it can be analysed. Thus, a child producing agent-action-locative (x + y) constructions should also produce both agent and action and locative constructions. But the converse need not be the case; the components do not guarantee the composition ... there is evidently, and this is simply an empirical discovery, some additional knowledge involved in putting the component items of knowledge together to make the more complex construction - in some cases, as shown in Stage II and also in the Brown and Hanlon paper, there may be quite long intervals, some months, between acquisition of the several component types of knowledge and their assemblage into the more complex construction" (p. 222).

Hypothesis III: There will be no difference in error scores among sentence types for groups or tasks.

The results found in Table 14 indicate a significant difference ($p = 0.003$) in error scores among sentence types. In addition, a significant ($p = 0.001$) interaction between tasks and sentence types was found. Therefore, Hypothesis III is rejected.

A discussion of the possible reasons for the task-sentence type interaction has been presented in the previous discussion of Hypothesis II. Of greatest interest in this section are the observed differences in error scores between sentence types for the comprehension task.

Consider Table 19 which contrasts the mean percent of errors scored for each sentence type found in Chapman and Miller (1975) and the present study.

Table 19. Mean Percent of Errors Scored on Sentence Types for Chapman and Miller (1975) and the Present Study

Sentence Type	Chapman-Miller (1975)	The Present Study
(1) Animate-animate	33.5	34.2
(2) Inanimate-inanimate	34.8	39.2
(3) Animate-inanimate	6.2	59.2
(4) Inanimate-animate	49.9	15.8

No appreciable difference in errors between the two studies is found for sentence types (1) and (2). However, the two studies differ greatly in the number of errors scored on sentence types (3) and (4). In fact, the present

study reverses the trend found in the Chapman and Miller (1975) study.

The finding that subjects scored the most errors on sentence type (4) and the least on sentence type (3) led Chapman and Miller (1975) to propose that children adopt the following strategy in determining subject-object assignment in the comprehension task:

> "assign agent status to the animate noun and object of action status to the inanimate noun" (p. 368).

These authors propose that many of the sentences heard by the young child are of the agent-action-object form, thus rendering their proposed strategy a plausible one. They concede, however, that such a lexical strategy can be overridden by event properties found in the context ie. the first object to move is assigned subject status.

Clearly, the proposed strategy is not supported by the present data. The children showed a marked preference for assigning agent status to the so called inanimate objects. Thus, the large percent of errors found for sentence type (3). How may these differences be explained?

Chapman and Miller (1975) reported that the objects used in the comprehension trials confounded size and color; that is, the animate toys were smaller and more colorful than the inanimate toys. The same holds true for the present study as the toys used here were proportionate in size. Hence, the observed differences between Chapman and Miller (1975) and the present study are not attributable to

such factors. A reasonable explanation may be that the semantic features defining agent were different for the children of each study. According to Chapman and Miller (1975), animacy was the relevant feature in determining a noun's assignment to agent status. However, it must be kept in mind that the defining characteristics of such semantic features as animate-inanimate may be different for the child and adult. The children of the present study appear to have considered all nouns as animate and thus capable of agent assignment. Comments by Bowerman (1974) elaborate this point:

> "In establishing a linguistically relevant semantic concept like agent, a child may initially attend to only one or two features rather than the entire set of features which define the concept for the adult; the features he attends to may or may not be critical for the adult. For example, the childish version of the adult agent concept might be defined only by the feature 'that which is capable of independent movement'. This would result in the child's treating of cars, machinery and the like as agents" (p. 204).

Since the children of the present study have considered all nouns as equally likely agents, another strategy had to be operable in assigning agent status to the various nouns in the comprehension task.

In her discussion of comprehension strategies during the stage of early preoperations (two to four years), Chapman (1977) proposes that children use lexical strategies as well as information gleaned from the context and past experience to determine sentence meaning. She defines the "probable event strategy" as one where the child interprets

a reference to events in terms of the usual relations that hold between them. That is, the child interprets a message in light of what she knows to be the usual relations between events and objects related in the sentence. This could explain the marked preference for assigning agent status to the so called inanimate nouns of car, truck and boat. The children in the present study adopted a probable event strategy based on the knowledge that "cars hit people; people don't hit cars" in determining subject-object assignment. Thus, the tendency to reverse the assignments for sentence type (3).

In sum, the proposed strategy for assigning agent status to animate nouns and object of action status to inanimate nouns, holds only if one assumes that the defining characteristics of each does not vary among children. This assumption is clearly not supported either in the literature (Bowerman, 1974) or by the empirical evidence of the present study.

Hypothesis IV: There will be no difference in the number of scorable responses between groups or tasks.

The results found in Table 18 indicate no significant difference ($p = 0.073$) between groups for the number of scorable responses obtained in the comprehension and production tasks. However, a significant difference ($p = 0.018$) in favor of comprehension was found between tasks. No significant interaction ($p = 0.073$) was found among groups

and tasks. Therefore, Hypothesis IV is rejected.

How might the observed differences in scorable responses between the two tasks be explained? If the task requirements favored production performance, the number of scorable responses for production should also exceed comprehension. This, however, is not the case as subjects were scorable on significantly greater numbers of trials in the comprehension task. Could this be evidence that the production task was in fact the more difficult?

In answering the above question a distinction must be made between attempting a task and subsequently performing that task. In the latter case, previous discussion has identified several factors which enhance production performance. The former distinction may be defined as the initial engagement of the child's cooperation in "playing the game" of each task. In this case, the comprehension task is favored.

One reason for this may be that the production task offered a wider range of possible responses, a factor which initially prompted Baird (1972) and Fernald (1972) to examine the early ICP studies. This is strongly related to the difficulty in explaining to the child the requirements of the production task without inadvertently modelling or training the desired response. In contrast, the comprehension task was easily demonstrated. The children readily and enthusiastically engaged in knocking the stimulus toys together. In fact, several subjects showed a marked pre-

ference for the comprehension task as evidenced by their requests that the production task be terminated or that they be allowed to demonstrate the actions themselves. Further, it would seem that the comprehension task involved a play routine that was more basic than the production task. The routine apparent in the comprehension task resembles what Piaget and Inhelder (1969) term "exercise play" in which an acquired behavior pattern is repeated for simple functional pleasure or for the enjoyment of causing an effect and of confirming an acquired skill.

The latter point raises a serious question in relation to the comprehension task. If the children engaged in a mere motor routine, it is possible that the verbal instruction was not construed by the child as a relevant factor in the performance of the task. In other words, was the child able to unwittingly perform the task without realizing and considering all relevant factors?

For the subjects of Group II, evidence to the contrary is found in the observation that these children exhibited behaviours indicative of their concern for the instruction during the testing session. For example, the children requested the repetition of instructions, confirmed instructions by such remarks as "this one" while holding up the subject toy and revised responses midway after an obvious contemplation of the test sentence.

Such was not the case for the subjects of Group I. These children did not display the above behaviours and

frequently attempted to respond before the instruction was issued. Further, they generally persisted with their first impulse even after frequent reminders to "listen and do what I tell you". These observations may indicate that subjects responded on the basis of a preconceived notion of what relations usually hold between the given objects. This is consistent with the finding that the greatest number of errors were scored on sentence type (3) and the least on sentence type (4). Further, the children were found to respond correctly to sentence types (1) and (2) an average of 57 and 58 percent, respectively, of trials. Thus, on those sentence types judged to be at least semantically confounding, the children performed at approximately a chance level.

These findings indicate that the children of Group I either ignored the verbal instruction entirely or were unable to use the word order cues found in the test sentences. In either case, it remains unclear whether the correct responses were obtained as a result of the child's understanding of the relation between subject and object that word order encodes.

CHAPTER SEVEN

CONCLUSIONS, LIMITATIONS AND IMPLICATIONS

The present study introduced four hypotheses in testing the predictions made by Chapman and Miller (1975) that (1) productive control of the subject-object relation exceeds comprehension performance based on syntactic form alone, and (2) less linguistically advanced children decode subject-object relations lacking referential support on the basis of a lexical-semantic strategy which assigns animate nouns to subject status and inanimate nouns to object status.

Hypothesis I which postulated that no difference in error scores would be found between groups for tasks or sentence types was supported. Younger, less linguistically advanced subjects appeared to perform as well as their older, linguistically superior counterparts on both tasks and all sentence types. This finding has been attributed to the relative conservatism of the statistical test used in the analysis and the small N of each group. It was also suggested that up to Stage IV (Brown, 1973), productive ability as measured by MLU may not be related to task performance.

Hypothesis II which stated that no difference in error scores would exist between the comprehension and production tasks was not supported. It was found that subjects scored significantly greater numbers of errors on the comprehension as compared to the production task. However, two confound-

ing variables were identified that may have enhanced production performance. First, the task requirements for comprehension and production appeared unequal since visual cues operative in the production task were not available during comprehension trials. Secondly, sentence types (3) and (4) may have been less semantically reversible than sentence types (1) and (2). This factor appeared to affect comprehension but not production performance as evidenced by the significant interaction effect among sentence types and tasks. In sum, these factors may have determined the observed differences in comprehension and production performance. Due to this, the results of the present study cannot be taken as evidence in support of the prediction made by Chapman and Miller (1975) that productive control of the ordering of subject and object exceeds comprehension performance based on syntactic form alone. Thus, the relationship between comprehension and production of syntaxtic form remains uncertain.

Hypothesis III proposed that no differences in error scores would be found among the four sentence types which were designed to test the subject-object relation. It was found that significant differences in error scores did exist among sentence types for the comprehension task. Significantly greater numbers of errors were scored on sentences which possessed an animate subject and inanimate object. Further, errors scored on sentences with an inanimate subject and animate object were significantly

less than all other sentence types. Hence, Hypothesis III was rejected.

These findings do not support the prediction made by Chapman and Miller (1975) that children decode subject-object relations on the basis of a lexical-semantic strategy which assigns subject status to the animate noun and object status to the inanimate noun. In fact, exactly the reverse trend was apparent in the present study since so called inanimate nouns were generally ascribed subject status and animate nouns, object status. The lexical-semantic strategy proposed by Chapman and Miller (1975) is based on the assumption that the semantic features of animate-inanimate which may define the concepts of agent and object are invariant among children. The literature (Bowerman, 1974) and the empirical evidence of the present study contradict this assumption. The children of the present study did not appear to make the animate-inanimate distinction that the subjects of Chapman and Miller (1975) displayed. In assigning subject-object status, these children appeared to adopt a "probable event" strategy in which the relations demonstrated were consistent with the relations between objects and events as they occur in the real world. Since no significant difference in error scores was found among sentence types for the production task, it would appear that different strategies were at work during these trials.

These facts support Bloom's (1978) contention that the different strategies employed in the performance of

comprehension and production tasks may vary as a function of the child's experience and cognitive-linguistic capacity. Further, the importance of the multi-linguistic and non-linguistic variables which comprise the communicative context in which messages are understood and produced is substantiated.

Hypothesis IV stated that no difference in scorable responses would be found between groups or tasks. While no significant group differences were obtained, the number of scorable responses on the comprehension task was found to be significantly greater than the production task.

Three factors account for the greater number of scorable responses obtained by subjects during the comprehension task. First, the production task offered a wider range of possible responses. Secondly, the response requirements for the production task were difficult to communicate to the child without modelling or inadvertently training the desired structure. Finally, the comprehension task appeared to be a more appealing activity as evidenced by the children's repeated requests to participate.

Evidence also indicated that the subjects of Group I may have responded to the comprehension items without attending to the test sentence issued. In effect, correct responding was at chance levels on sentence types (1) and (2) which were considered to be semantically neutral items.

This implies that the comprehension score obtained for each subject was not a valid indicator of the child's ability to use word order as a comprehension strategy in determining subject-object assignment. Thus, it is questionable whether subjects of Group I used word order as a cue during the comprehension task.

In conclusion, the predictions made by Chapman and Miller (1975) have not received unqualified support here. The proposed lexical-semantic strategy for comprehension was not employed by the subjects of the present study; these children appeared to use a probably event strategy in determining subject-object assignment. Although productive control of the ordering of subject and object was found to exceed comprehension performance, the facilitating effect of visual cues available during the production trials likely contributed to the observed differences in performance between the two tasks.

These conclusions must be considered in light of the limitations of the present study. First, as in other studies of language development, the sample size of the present study was small (N = 10). In addition, children at this age display considerable variation in the development of language skills. These factors limit the degree to which the findings of the present study can be generalized to the general population. Secondly, the stimulus objects used in testing the subject-object relation were not of uniform size. That the "animate" toys were smaller than the

"inanimate" toys may have influenced comprehension performance. Thirdly, sentence types (3) and (4) appeared to be less semantically reversible than sentence types (1) and (2); thus, the role of word order in the subjects' comprehension was obscured by their use of semantic strategies. Thirdly, the task requirements for comprehension and production may have been unequal since visual cues operative in the production task were not available to the child during the comprehension trials. In effect, the children were required to rely on memory more during the comprehension than production task. Finally, it is not clear what relation the child's performance on these contrived tasks may have to her comprehension or productive ability in a natural situation.

These limitations provide direction for improving the methodology used in future studies which may address the question of the relation between comprehension and production of sentence structure. Larger sample sizes should be employed. The data obtained from constructed tasks should be compared to the child's performance in her natural setting. Task requirements should be carefully regulated. For example, in the present study, the task requirements would have been better controlled if the toys had been moved back to their original positions after each production trial demonstration. In this way, the child would have had to rely on memory as much for the production as the comprehen-

sion task. However, this assumes that a visual presentation is equally as memorable as an auditory one which in itself requires verification.

As Ingram (1974) has stated, it can be argued *ad nauseum* whether one constructed task is more difficult than another. However, it may well be the case as Bloom (1978) and others propose, that the tasks simply cannot be equated since comprehension and production appear to be two mutually dependent but separate processes each with its own intervening variables. As such, we can expect no linear or constant relationship between the two processes to emerge.

Perhaps, as Ingram (1974) has indicated, the real issue is not whether one process precedes the other; the important questions that remain involve the nature of each process and the changes in their relationship that occur in the course of development. It is in these areas that the most fruitful research will take place.

BIBLIOGRAPHY

1. Abramson, A. Child Language: An Interdisciplinary Guide to Theory and Research. Baltimore: University Park Press, 1977.
2. Baird, R. On the Role of Chance in Imitation, Comprehension and Production: Test Results. Journal of Verbal Behavior, 1973, 11, 474-477.
3. Bates, E. Language and Context: The Acquisition of Pragmatics. New York: Academic Press, 1976.
4. Bloom, L. Language Development: Form and Function in Emerging Grammars. Cambridge, Mass.: MIT Press, 1970.
5. Bloom, L. Talking, Understanding and Thinking. In R. Schiefelbusch and L. Lloyd (Eds.), Language Perspectives: Acquisition, Retardation and Intervention. Baltimore: University Park Press, 1974.
6. Bloom, L. Developmental Interaction Between Understanding and Speaking. In L. Bloom and M. Lahey, Language Development and Language Disorders. New York: John Wiley and Sons, Inc., 1978.
7. Bowerman, M. Early Syntactic Development: A Cross-Linguistic Study with Special Reference to Finnish. Cambridge, England: Cambridge University Press, 1973.
8. Braine, M. The Ontogeny of English Phrase Structure: The First Phase. Language, 1963, 39, 1-14.
9. Brown, R. A First Language: The Early Stages. Cambridge, Mass: Harvard University Press, 1973.

10. Brown, R., and Fraser, C. The Acquisition of Syntax. in C. Cofer and B. Musgrave (Eds.), Verbal Behavior and Learning: Problems and Processes. New York: McGraw-Hill, 1963.

11. Brown, R., and Hanlon, C. Derivational Complexity and Order of Acquisition in Child Speech. In J.K. Hayes (Ed.), Cognition and the Development of Language. New York: Wiley, 1970, 155-207.

12. Chapman, R. Comprehension Strategies in Two and Three Year Olds: Animate Agents or Probable Events? Journal of Speech and Hearing Research, 1978, 21, 746-761.

13. Chapman, R., and Miller, J. Word Order in Early Two and Three Word Utterances: Does Production Precede Comprehension? Journal of Speech and Hearing Research, 1975, 18, 355-71.

14. Clark, R., Hutcheson, S., and Van Buren, P. Comprehension and Production in Language Acquisition. Journal of Linguistics, 1974, 10, 39-54.

15. Cohen, M. Active and Passive Language Development of Grammar. Connecticut College of Psychology Journal, 1967, 4, 20-24.

16. DeLaguna, G. Speech: Its Function and Development. New Haven: Yale University Press, 1927.

17. DeVilliers, J., and DeVilliers, P. Development of the Use of Word Order in Comprehension. Journal of Psycholinguistic Research, 1973, 2, 331-341.

18. Dore, J. Holophrases, Speech Acts and Language Universals. Journal of Child Language, 1975, 2, 21-40.

19. Fernald, C. Control of Grammar in Imitation, Comprehension and Production: Problems of Replication. Journal of Verbal Learning and Verbal Behavior, 1972, 11, 606-613.

20. Fillmore, C. The Case for Case. In E. Bach and R. Harms (Eds.), Universals in Linguistic Theory. New York: Holt, Rhinehart and Winston, 1968, 1-87.

21. Fraser, C., Bellugi, U., and Brown, R. Control of Grammar in Imitation, Comprehension and Production. Journal of Verbal Learning and Verbal Behavior, 1963, 2, 121-135.

22. Guillaume, P. Les debuts de la phrase dan le language de l'enfant. Journal of Psychology, 1927, 24, 1-25.

23. Greenfield, P., and Smith, J. The Structure of Communication in Early Language Development. New York: Academic Press, 1976.

24. Halliday, M. Learning How to Mean. In E. Lenneberg and E. Lenneberg (Eds.), Foundations of Language Development: A Multidisciplinary Appraoch. New York: Academic Press, 1975.

25. Huttenlocker, J. The Origins of Language Comprehension. In R. Solso (Ed.), Theories in Cognitive Psychology. New York: Halstead, 1974.

26. Huttenlocker, J., Eisenberg, K., and Strauss, S. Comprehension: Relation Between Perceived Actor and Logical Subject. Journal of Verbal Learning and Verbal Behavior, 1968a, 1, 527-530.

27. Huttenlocker, J., and Strauss, S. Comprehension and a Statement's Relation to the Situation it Describes. Journal of Verbal Learning and Verbal Behavior, 1968b, 7, 300-304.

28. Huttenlocker, J., and Weiner, S. Comprehension of Instructions in Varying Contexts. Cognitive Psychology, 1971, 2, 369-385.

29. Ingram, D. The Relationship Between Comprehension and Production. In R. Schiefelbusch and L. Lloyd (Eds.), Language Perspectives: Acquisition, Retardation and Intervention. Baltimore: University Park Press, 1974.

30. Keeny, T., and Wolfe, J. The Acquisition of Agreement in English. Journal of Verbal Learning and Verbal Behavior, 1972, 11, 698-705.

31. Lahey, M. The Use of Prosody and Syntactic Markers in Children's Comprehension of Spoken Sentences. Journal of Speech and Hearing Research, 1974, 17, 656-668.

32. Leopold, W. Speech Development of a Bilingual Child. Illinois: Northwestern University Park Press, 1939-49.

33. Lovell, K., and Dixon, E. The Growth of Control of Grammar in Imitation, Comprehension and Production. Journal of Child Psychological Psychiatry, 1967, 8, 31-39.

34. McCarthy, D. Language Development in Children. In L. Carmichael (Ed.), A Manual of Child Psychology. New York: Wiley, 1954, 492-630.

35. Miller, W., and Ervin, S. The Development of Grammar in Child Language. In U. Bellugi and R. Brown (Eds.), The Acquisition of Language. Monographs of the Society for Research in Child Development, 1964, 29, 9-34.

36. Nurss, J., and Day, D. Imitation, Comprehension and Production of Grammatic Structure. Journal of Verbal Learning and Verbal Behavior, 1971, 10, 68-74.

37. Petretic, P., and Tweney, R. Does Comprehension Precede Production: the Development of Children's Responses to Telegraphic Sentences of Varying Grammatic Adequacy. Journal of Child Language, 1977, 4, 201-209.

38. Schlesinger, I. Production of Utterances and Language Acquisition. In D. Slobin (Ed.), The Ontogenesis of Grammar. New York: Academic Press, 1971, 63-101.

39. Shipley, E., Smith, C., and Gleitman, L. A Study in the Acquisition of Language: Free Responses to Commands. Language, 1969, 45, 322-343.

40. Slobin, D. Universals of Grammatic Development in Children. In G.B. Flores D'Arcais and W.J. Levelt (Eds.), Advances in Psycholinguistics. Amsterdam: North-Holland Publishing Co., 1970.

41. Stern, W. The Chief Periods of Further Speech Development in Psychology of Early Childhood. (A. Berwell, Trans.) London: Allen and Unwin, 1930.

42. Turner, E., and Rommetveit, R. The Acquisition of Sentence Voice and Reversibility. Child Development, 1967, 38, 649-660.

43. Wetstone, H., and Friedlander, B. The Effect of Word Order on Young Children's Responses to Simple Questions and Commands. Child Development, 1973, 44, 734-740.

44. Winer, B. Statistical Principles in Experimental Design. New York: McGraw-Hill, 1971.